TI 2026-069/IV
Tinbergen Institute Discussion Paper

Blockchains and Strategic Secondary Censorship

Yackolley Amoussou-Guenou [1]
Maarten R.C. van Oordt [2]

1 Université Paris-Panthéon-Assas, CRED

2 Vrije Universiteit Amsterdam and Tinbergen Institute

Blockchains and Strategic Secondary Censorship*

Yackolley Amoussou-Guenou†
Université Paris-Panthéon-Assas, CRED

Maarten R.C. van Oordt‡
Vrije Universiteit Amsterdam
Tinbergen Institute

September 1, 2026

Abstract

Censorship resistance is often considered to be a core attribute of distributed ledgers. Censorship resistance refers to the inability to selectively exclude technically valid but undesirable transactions from the blockchain. We examine blockchain censorship in a game-theoretic framework that allows for both primary and secondary censorship. The analysis identifies scenarios in which both inclusion and censorship equilibria exist. Once an equilibrium with strategic secondary censorship is implemented, it may be hard to revert to inclusion: Censorship equilibria are perfectly coalition-proof if the negative impact of an undesirable transaction on block producers is sufficiently large. These results highlight the growing importance of research into methods shaping censorship resistance at the technical layer.

Keywords: blockchain, consensus, cryptocurrency, distributed ledger, game theory.
JEL Codes: C72, G2, L86

*We are thankful to Jonathan Chiu, Cyril Monnet, Charles M. Kahn, and participants in the ANR MonFinTech Kick-off Workshop (2026) and seminars at the Bank of Canada (2026) and Vrije Universiteit Amsterdam (2026) for helpful comments and suggestions. The first author gratefully acknowledges the Agence Nationale de la Recherche for financial support (Grant ANR-25-CE26-3025, MonFinTech).

†Email: yackolley.amoussou-guenou@assas-universite.fr
‡Email: m.van.oordt@vu.nl

I. Introduction

From community debates over freezing quantum-vulnerable Bitcoin addresses to sanction-induced blacklisting of addresses by some participants in the Ethereum network, major blockchains are increasingly exposed to attempts to exclude certain transactions. These developments draw renewed attention to the concept of censorship resistance, which refers to the inability to selectively exclude technically valid transactions. Censorship resistance is often considered to be a core attribute of distributed ledgers. The standard rationale typically proceeds as follows: "With a distributed ledger, many independent block producers can include a particular transaction in a block. If at least some block producers are willing to include that transaction, it will eventually be included in the ledger. Even if many block producers were to censor the transaction, the worst consequence would be a delay before its inclusion."

This standard rationale for explaining the censorship resistance of distributed ledgers focuses solely on what we define as *primary censorship*. Primary censorship in the context of distributed ledgers is the choice of block producers to selectively exclude technically valid but undesirable transactions from their own blocks.[1] It differs from what we define as *secondary censorship*. Secondary censorship is the choice to selectively ignore blocks produced by others if those blocks contain or confirm technically valid but undesirable transactions. Primary censorship targets the content a block producer includes in her own block, whereas secondary censorship targets the content of blocks produced by third parties.[2]

The theoretical possibility of secondary censorship is difficult to eliminate for widely used consensus protocols. For proof-of-work protocols, such as the one used by Bitcoin, block building occurs randomly, and block producers can collectively act as if they never observed

[1] If primary censorship were the only possible form of censorship, then a single block producer who is willing to include an undesirable transaction could be sufficient to prevent the censorship of that transaction.

[2] We borrow the labels *primary* and *secondary* from the terminology used in economic sanctions: *primary sanctions* restrict a country's own dealings with a sanctioned foreign entity; *secondary sanctions* target the dealings of foreign third parties with those subject to primary sanctions (e.g., Meyer, 2009; Morgan et al., 2023).

a broadcast block whose contents include an undesirable transaction. Ignoring blocks is also possible under consensus protocols where block producers broadcast a new block during their designated slots, such as proof-of-stake or proof-of-authority. These protocols ensure *liveness* by including mechanisms to handle inactivity. Such mechanisms prevent a blockchain from halting when a block producer appears to be offline. At the same time, these mechanisms enable others to collectively disregard a block producer's actions by effectively treating them as if they were offline.

Given the technical possibility, a key question concerns the conditions under which censorship of transactions may emerge on a blockchain. This paper shows that interesting strategic interactions arise when accounting for secondary censorship decisions by block producers in a game-theoretic framework. Whether inclusion or censorship emerges critically depends on the anticipated actions by other block producers. Strategic secondary censorship becomes reinforcing if including a transaction has a sufficiently large negative impact on other block producers. The threat of secondary censorship may then prompt all block producers to engage in primary censorship, even for transactions that offer a substantial fee. Strategic secondary censorship is never a guaranteed outcome, because inclusion remains optimal if other block producers continue to include the transaction. Censorship equilibria may be difficult to coordinate on, but once block producers collectively implement strategic secondary censorship, such censorship may become pervasive: block producers would not profit from switching back to an inclusion equilibrium if the inclusion of the transaction has a sufficiently large negative impact on them.

We study an environment in which block producers may censor a transaction that is undesirable in the sense that its inclusion would negatively affect all block producers.[3] Block producers can choose whether to include the undesirable transaction in their blocks, and they may also choose the specific prior block on which to build their new block (e.g., acting as if the producers of ignored blocks were offline). The block producer who chooses to include the

[3]We provide a variety of examples of such transactions in Section II.

undesirable transaction may earn a larger block reward because of an incremental transaction fee, but she receives block rewards only if her block becomes part of the consensus chain. The model also accounts for the fact that block producers' payoffs may be negatively affected by a failure to reach consensus promptly, which translates into empty slots or orphaned branches.

This game-theoretic framework highlights several strategic considerations that optimizing block producers face. A block producer who decides whether to include an undesirable transaction must consider not only the payoff from inclusion conditional on her block becoming part of the consensus chain but also whether inclusion affects other producers' willingness to build on her block. Block producers need to assess whether a threat of secondary censorship exists and, if so, whether the threat is credible. Another strategic consideration is the anticipated inclusion of transactions by subsequent block producers. If a transaction with negative consequences will enter the consensus chain regardless, then the current block producer may as well include it immediately to collect the incremental transaction fee.

We use an extension of our model to analyze whether a jurisdiction could induce the censorship of a particular transaction by imposing a penalty on the producer who includes that transaction in her block. The results demonstrate that penalties can incentivize domestic block producers to apply primary censorship, but that they are insufficient to eliminate inclusion equilibria unless the jurisdiction has enforcement powers over an overwhelming share of block producers. The difficulty lies in the fact that foreign block producers are incentivized by transaction fees to include the transaction, while domestic block producers prefer to avoid failures to reach consensus. With limited enforcement powers, penalties are anticipated to transform inclusion equilibria into equilibria with delayed inclusion. This theoretical result is in line with empirical evidence suggesting that sanctions imposed by the U.S. Office of Foreign Assets Control (OFAC) have induced primary censorship and thereby delayed the processing of certain transactions on the Ethereum blockchain (Brownworth et al., 2024; Wahrstätter et al., 2024). Penalties rule out inclusion equilibria only if the jurisdiction has

enforcement powers over an overwhelming share of block producers, and even then only when the targeted transaction imposes a sufficiently large loss on all block producers.

The remainder of this paper is organized as follows. Section II provides various examples of transactions that could be targeted by censorship in our model. Section III discusses related literature. Section IV introduces the model. Section V characterizes the equilibria if block producers consider primary censorship but ignore the possibility of secondary censorship, as is sometimes implicitly assumed in popular discourse. Section VI characterizes the equilibria if block producers consider strategic secondary censorship. Section VII extends the model by introducing a jurisdiction that seeks to induce censorship by penalizing block producers who include a specific transaction in their block. Section VIII discusses model modifications such as side-payments to promote inclusion and block producers who propose multiple consecutive blocks. Appendix A uses a simple game to illustrate the difference between the two equilibrium concepts employed in this paper: Subgame Perfect Equilibrium and Perfectly Coalition-Proof Nash Equilibrium. Proofs are in Appendix B.

II. Background

Our model considers censorship outcomes for transactions that are undesirable in the sense that they negatively affect all block producers' payoffs. Our framework is agnostic about the precise source of a transaction's negative impact, which may be pecuniary or not. The negative impact may stem from within the cryptocurrency ecosystem or from external factors. Several scenarios fit this general description. Some lie at the centre of ongoing debates within the cryptocurrency realm, others pertain to historical episodes, and still others remain hypothetical. We provide five scenarios below.

Scenario 1: Post-Quantum Migration of Bitcoin

As quantum computing continues to advance, a central debate within the Bitcoin community concerns whether transactions involving potentially vulnerable Bitcoin addresses should ultimately be censored. This could affect an estimated 5.6 million dormant bitcoins, more

than 25 percent of the total supply. Those bitcoins will become vulnerable to theft once quantum computing advances to the point at which the cryptographic keys for early and reused Bitcoin addresses can be broken (Acuna, 2026a).[4] The vulnerable addresses include those that were created in the early years of Bitcoin's history.

One proposed path forward is to censor transactions from vulnerable addresses unless the owners claim ownership of dormant bitcoins using quantum-resistant cryptography within a given time window. Opponents of this idea argue that freezing addresses would be a departure from core principles, whereas proponents emphasize the risk of a market panic if vulnerable bitcoins were to flood the market (Acuna, 2026b). A negative price reduces the present value of future mining rewards, which imposes a loss on block producers. Lopp et al. (2026) specifically point to such a loss in their formal proposal (BIP-361) when they write that "[a] quantum attack on Bitcoin will significantly devalue both their [mining] hardware and Bitcoin as a whole."

Viewed through the lens of our model, the ongoing debate on whether or not to ultimately freeze vulnerable addresses could be seen as a coordination effort by the Bitcoin community to select either an equilibrium in which transactions from vulnerable addresses are permitted or one in which such transactions are censored. Our model shows that both equilibria are possible, with the censorship equilibrium being coalition-proof if the anticipated loss to block producers from the theft of vulnerable bitcoins is sufficiently large.

Scenario 2: Transactions Involving Code Exploits

Other transactions to which our model applies are transactions arising from code exploits. Although such transactions may be technically valid under the protocol's rules, they produce

[4]It is expected that quantum computers will advance to a level where ECDSA private keys and Schnorr signatures can be derived economically from exposed public keys (Acuna, 2026b). Vulnerable addresses include those based on a legacy format that exposes the public key when an address is created (P2PK). These formats were used during Bitcoin's earliest years. Certain other address formats conceal the public keys but reveal the public key once their owner spends from them (e.g., P2PKH). Bitcoins held in this second category of addresses become vulnerable to quantum attacks if their owner reuses the address after spending from it (Dragos et al., 2020; Barmes et al., 2020; Bertucci et al., 2025).

unintended outcomes. Decisions to censor transactions related to code exploits have been heavily debated and have produced different outcomes in different contexts.

A prominent example of such a debate within the Ethereum community occurred following the attack on the so-called DAO smart contract in June 2016. At the time of the attack, roughly one-seventh of all existing ETH had been committed to the DAO smart contract. An adversary exploited a vulnerability in this smart contract to drain over 3 million ETH from it into a smart contract under the adversary's control. Because the siphoned funds could not be moved for 28 days, the Ethereum community debated whether to censor future transactions from the adversary's address to prevent the adversary from escaping with the funds (e.g., Buterin, 2016). The final outcome of the debate was an extraordinary intervention that altered account balances on the Ethereum blockchain to reverse the consequences of the adversary's actions as documented in EIP-779 by Detrio (2017).

Actions to censor and revert technically valid transactions that involve code exploits or hacks are not uncontested. Those opposing the intervention following the attack on the DAO smart contract continued the original chain, resulting in a competing chain known as Ethereum Classic. Censorship or reversion of transactions seems generally the exception rather than the rule. In the years since, various hacks and smart-contract exploits have resulted in substantial losses without triggering similar interventions. Even a proposal to roll back transactions linked to the theft of an exceptional sum of roughly USD 1.4 billion from the Bybit cryptocurrency exchange reportedly met with community ire (Godbole, 2025).

Scenario 3: Pressure from a Major Jurisdiction

One hypothetical scenario to which our model applies is one in which a major jurisdiction seeks to exert influence on which transactions can be included in the blockchain. Governments could implement a policy that prohibits local trading in the native token of a blockchain if the chain processes certain controversial transactions or transactions involving blacklisted addresses. The inclusion of those transactions could then harm block producers because delistings of tokens from centralized exchanges are generally associated with price

declines (e.g., Reback and Crawley, 2024). The price impact of delistings is expected to be larger if the threat were to originate from a major jurisdiction or if multiple jurisdictions were to coordinate their actions. The threat of price declines could incentivize block producers to coordinate on a censorship equilibrium to avoid a reduction in the value of staked tokens or future mining rewards.

Scenario 4: Guild Attacks

Another hypothetical scenario to which our model applies is what we call a *guild attack*. Under proof-of-stake, entry requires existing stakers to include the transaction that admit the new stakers. Incumbent stakers could form a guild that systematically censors such transactions with the goal of increasing the value of membership for the incumbents. The value of membership would be diluted by any transaction that expands the set of members, thereby imposing a loss on existing stakers.

Scenario 5: Soft Forks

Our model can also be interpreted in the context of certain types of *soft forks*. Narayanan et al. (2016, p. 73) describe those as updates "*...to add features that make validation rules stricter. That is, they restrict the set of valid transactions (or the set of valid blocks) such that the old version would accept all the blocks, whereas the new version would reject some of the blocks accepted by the old version.*" Such soft forks are not ordinarily described in terms of "censorship", but they are similar insofar as they entail the deliberate exclusion of transactions that were previously regarded as technically valid.

Soft forks generally implement improvements, such as changes designed to reduce protocol vulnerabilities. When such an improvement is expected to enhance the network's value, then the continued inclusion of transactions permitted under the old rules imposes an implicit loss on block producers insofar as that inclusion prevents the anticipated increase in network value. These updates are usually carried out via standardized procedures. In the context of

our model, these standardized procedures can facilitate coordination among participants on a "censorship equilibrium" that implements the soft fork.

III. Related Literature

The present paper develops a game-theoretic framework to assess whether block producers may strategically apply secondary censorship to exclude transactions that affect them negatively. Only a small number of other papers use a game-theoretic approach to study blockchain censorship, and those examine aspects other than strategic secondary censorship. Kiayias et al. (2021) analyze strategic interactions between the manager of a staking pool and the contributors to the pool in an environment that permits primary censorship. They demonstrate that primary censorship by pool managers may result in the delayed inclusion of transactions, despite the preference of contributors for the inclusion of all transactions. Berger et al. (2025) examine whether a third party can bribe block producers to apply primary censorship to a transaction—an issue that is particularly salient in the context of programmable payments, where a third party may seek to prevent the release of funds to the payee.

Our results reveal that the possibility of strategic secondary censorship may enable outcomes where certain transactions are persistently excluded from the blockchain. These results suggest a growing need for game-theoretic research into new mechanisms that aim to strengthen censorship resistance at the technical layer. Proposed mechanisms include *inclusion lists*, under which multiple validators are empowered to require the current block proposer to include specific transactions (Thiery et al., 2024; Wadhwa et al., 2025), as well as the introduction of *multiple concurrent proposers* (Garimidi et al., 2025). A planned Ethereum upgrade is set to promote censorship resistance by introducing inclusion lists through the implementation of the FOCIL proposal (EIP-7805) of Thiery et al. (2024). Other proposed avenues for enhancing censorship resistance include increasing the diversity of the validator

set (John et al., 2025a), techniques such as *transaction ordering* (Cachin et al., 2022; Gramoli et al., 2024), and the impact assessment of past protocol updates such as *proposer-builder separation* (Heimbach et al., 2023). Some analyses of censorship resistance are based on the premise that a large share of the network will follow protocol rules that encourage inclusion. Understanding the extent to which protocol enhancements can promote censorship resistance in a game-theoretic environment in which block producers behave strategically and consider secondary censorship will be an important avenue for future research.

Our paper contributes to a growing literature that applies a game-theoretic approach to study the behaviour of nodes that extend distributed ledgers. Much of that work provides valuable, complementary insights into areas beyond blockchain censorship. Biais et al. (2019) analyze consensus in an environment that permits blockchain forks and establish that Bitcoin's longest-chain rule is an equilibrium, even though multiple equilibria exist. Hałaburda et al. (2021) analyze transaction propagation in networks with malicious nodes and derive the minimum reward needed for an equilibrium in which the network reaches consensus on the acceptance of a transaction. Amoussou-Guenou et al. (2024) study outcomes for committee-based consensus protocols and demonstrate that members must be pivotal to ensure the validity of blocks and the liveness of the blockchain. Game-theoretic approaches have also devoted substantial attention to the extent to which consensus protocols can prevent attacks that attempt to double-spend or halt a network. Prominent examples include Chiu and Koeppl (2022), Garratt and Van Oordt (2023), Budish (2025), John et al. (2025b), and Ebrahimi et al. (2026). An excellent guide to further references and other modelling approaches is provided by Hałaburda et al. (2022).

IV. Model

We model the process of writing the blockchain as a multi-stage game with perfect information and an infinite number of stages. The blockchain provides a slot for a block at each

stage. At every stage indexed by an integer $t \geq 1$, a single player proposes the content for the block in slot t. Other players have no actions. Since there is only a single active player at every stage, we use subscript t to refer to both the stage and the active player, whom we refer to as block producer t.[5] The number of block producers is infinite. Every block producer gets a single opportunity to propose a block.[6]

A. Actions

Block producers readily include transactions that they regard as neutral. A block producer's action concerns her choices regarding the content of her block in two respects: (1) the reference to the most recent block that she accepts, and (2) whether she includes a certain undesirable transaction that negatively affects the payoffs of all block producers.[7] The twofold choice of block producer t is denoted by the tuple $a_t = (k_t, e_t)$. We will impose restrictions on the permissible action set to compare equilibria in different scenarios (e.g., with and without strategic secondary censorship).

By choosing k_t, block producer t indicates the slot of the last block she accepts.[8] We use the convention that k_t denotes the number of slots that should be ignored to find the last valid block the producer accepts. If the current block producer accepts the block proposed for the previous slot, then she sets $k_t = 0$. If the current block producer ignores the previous slot but accepts the block in the slot before, then she sets $k_t = 1$, and so forth. Block

[5] We opt for the generic term *block producer* to refer to the agent with the final say on the content of a proposed block. With proof-of-work, block producers are commonly referred to as miners. With some proof-of-stake protocols, block producers are referred to as block proposers.

[6] The one-block-per-producer structure prevents a small group of high-share block producers from swinging the blockchain towards censorship. Nevertheless, censorship equilibria will still emerge.

[7] The transaction is undesirable in the sense that all block producers incur a loss if the transaction is included in the consensus chain, so that its inclusion may provoke disapproval from other block producers. The loss may be pecuniary or non-pecuniary. Examples of pecuniary losses are the decline in the value of the stakes or mining hardware. Non-pecuniary losses may arise from the disutility or reputation cost of facilitating transactions for repugnant activities (Garratt, 2015). We do not restrict the term undesirable to the meaning of provoking *public* or *regulatory* disapproval, although it is not unthinkable that such disapproval could be the source of the loss that the transaction imposes on block producers.

[8] A block producer does so in practice by including the *hash of the parent block* in her block (i.e., a digital reference to the last block she accepts).

producer t can choose any integer value $k_t \in [\![0, t-1]\!]$. If $k_t = t-1$, then the current producer ignores all blocks that were proposed since the initial state of the blockchain.

Every block points back to a previous block. Hence, by accepting a previous block, the block producer implicitly accepts a chain of blocks. This blockchain can be constructed from the observable history of all actions, which is denoted by $h_t = (a_{t-1}, a_{t-2}, ...)$. Given block producer t's choice of k_t, the chain she accepts, including her own block and the initial state, consists of the set of blocks in slots $T(h_{t+1}) = \{T_0, T_1, ..., T_n\}$, where, $T_0 = t$, $T_1 = T_0 - 1 - k_{T_0}$, $T_2 = T_1 - 1 - k_{T_1}$, ..., $T_n = T_{n-1} - 1 - k_{T_{n-1}} = 0$. The number of empty slots in the chain accepted by the current block producer follows from $T(h_{t+1})$ as $k(h_{t+1}) = t - |T(h_{t+1})|$ or, equivalently, from summing the number of rejected blocks in the proposed chain $k(h_{t+1}) = \sum_{s \in T(h_{t+1})} k_s$.

The choice of $e_t \in \{0, 1\}$ indicates whether the block producer includes the undesirable transaction in her block ($e_t = 1$), or not ($e_t = 0$). A transaction can be included in a chain only once. The block producer can only include the undesirable transaction if the chain referenced by her block does not already include it. That is, if her choice for k_t is such that $e(h_{t-k_t}) = \sum_{s \in T(h_{t-k_t})} e_s = 0$. Otherwise, the undesirable transaction is already recorded on the chain she accepts, so that her choice will be constrained to $e_t = 0$.

B. Payoffs

Block producers face a delay of q blocks before they receive their payoffs, where q is some large constant. The payoff of block producer t materializes at time $t+q$; the payoff of block producer $t+1$ materializes at time $t+q+1$; and so forth. Time $t+q$ can be interpreted as the point at which block producer t is allowed to spend block rewards or to exit and sell her stake.[9]

[9]The value of q varies per consensus protocol but tends to be large in practice. Block rewards for Bitcoin become spendable only after 100 subsequent blocks have been mined (Bitcoin Core Developers, 2026a). Withdrawal times for Ethereum validators vary depending on the exit queue, and funds become eligible for withdrawals only after more than 256 epochs of 32 slots have passed since a validator's exit (Ethereum Foundation, 2025).

The payoff of a block producer depends on whether her block becomes an element of the *consensus* chain. The consensus chain comprises the blocks in the "longest" chain, where the longest chain is defined as the chain with the most blocks (and, hence, the smallest number of ignored blocks). At time t, the last block in the longest chain is defined as $t^*(t)$, where $t^*(t) = \min\left(\arg\max_{s\in[\![0,t-1]\!]} |T(h_{s+1})|\right)$.[10] We denote the set of blocks in the consensus chain given the history of all moves until time $t+q$ as $T^{\mathcal{C}}(h_{t+q}) = T(h_{t^*(t+q)})$ and the total number of rejected blocks in the consensus chain as $k^{\mathcal{C}}(h_{t+q}) = k(h_{t^*(t+q)})$.

Block producer t receives a benefit b if her block is an element of the consensus chain at time $t+q$. The parameter b captures block rewards such as newly minted coins and fees for transactions that are considered neutral among block producers.

A block producer may earn an incremental transaction fee f (or bribe) for including a transaction that is regarded as undesirable among block producers. The transaction is undesirable in the sense that all block producers incur a loss ℓ if the transaction is included in the consensus chain when their payoffs materialize, so that its inclusion may provoke disapproval from other block producers. Whether the undesirable transaction is included in the consensus chain by time $t+q$ is indicated by $e^{\mathcal{C}}(h_{t+q}) = e(h_{t^*(t+q)})$. The block producer earns the incremental fee f only if her block with the undesirable transaction is an element of the consensus chain at time $t+q$.

The final factor affecting the payoffs of block producers is the loss in the value of their stakes from interruptions in the consensus chain. Lack of consensus can cause transaction delays as well as uncertainty about whether operations in a proposed block will indeed become part of the consensus chain. We measure the loss from interruptions by the loss function $\delta(k^{\mathcal{C}}(h_{t+q}))$. This function is assumed to be a strictly increasing and concave function of the total number of ignored blocks, $k^{\mathcal{C}}(h_{t+q})$. This functional shape is intended to reflect the notion that a few missing blocks may have a limited negative impact on the value of the

[10] As a tie-breaking rule, we introduce the convention to define the longest chain as the "oldest" longest chain which explains the min-operator in the definition of the longest chain. This convention is consistent with the "first-seen" rule in the Bitcoin protocol that gives priority to the first observed branch if a node observes multiple branches with the same length (e.g., Bitcoin Core Developers, 2026b).

stakes, but that the negative impact increases if interruptions last longer. The concavity reflects the notion that there is an upper bound to the losses of block producers (e.g., they may lose at most the total value of their stakes). Moreover, the loss is assumed to be zero if all slots are filled (i.e., $\delta(0) = 0$).

Combining these components, we have that the payoff function of each block producer t is given by

$$u_t(k_t, e_t, h_{t+q}) = (b + e_t f)\mathbf{1}(t \in T^{\mathcal{C}}(h_{t+q})) - \delta(k^{\mathcal{C}}(h_{t+q})) - e^{\mathcal{C}}(h_{t+q})\ell, \tag{1}$$

where the indicator function $\mathbf{1}(t \in T^{\mathcal{C}}(h_{t+q}))$ equals one if block t is included in the consensus chain at $t+q$, and zero otherwise. When unambiguous, we suppress the dependence on h_{t+q} and write $u_t(k_t, e_t)$. Finally, we introduce two tie-breaking conventions in situations with equal payoffs in equation (1). First, we assume that block producers, when payoffs are equal, prefer the outcome in which the undesirable transaction is included. The sole reason for this convention is to avoid an otherwise obligatory discussion of boundary cases in which parameter values coincide. Second, we assume that block producers prefer to build upon an earlier block if they have to choose between two competing chains with equal payoffs.[11]

C. Equilibrium Concepts

The creation of the blockchain is a multi-stage game with observed actions. We analyze pure equilibria and use the equilibrium concept of subgame perfection to model the anticipated behaviour of block producers. The payoff of any block producer t in the model is unaffected by actions of block producers in the distant future, i.e., the actions of those playing beyond time $t+q$. This implies that payoffs are continuous at infinity. Hence, a *Subgame Perfect Equilibrium* (SPE) in our model can be proven using the one-stage deviation principle (Fudenberg and Tirole, 1991, Theorem 4.2).

[11] This convention becomes relevant in a subset of scenarios discussed in Section VI, in which multiple block producers have incentives to compete for a high transaction fee.

The model will admit multiple SPEs in some situations. In such SPEs, no individual block producer can profitably deviate, but it may be possible for coalitions of block producers to deviate profitably. In situations where there are multiple SPEs, we also assess whether an equilibrium constitutes a *Perfectly Coalition-Proof Nash Equilibrium* (PCPNE) in the sense of Bernheim et al. (1987, Section IV) and Peleg (1992).[12] An SPE is considered to be perfectly coalition-proof if there is no permitted deviation by any coalition of players that improves the payoffs of all players in that coalition. A deviation of a coalition is permitted only if the deviation itself is perfectly coalition-proof for the deviating coalition at any remaining stage (given the actions in previous stages). That is, if the coalition were to play the deviation, then, for any remaining stage, there must be no permitted deviations for any subgroup of players in the coalition that could further improve the payoffs of all players in that subgroup.[13] In the context of the present model, the PCPNE refinement effectively eliminates SPEs that depend on an non-credible threat by a coalition of subsequent block producers, in a similar way that backward induction can eliminate Nash equilibria that are supported by an non-credible threat from a single player in a sequential game. Appendix A provides an illustration of a PCPNE in a simple game with multiple SPE outcomes.

D. Some Benchmark Strategies

For expositional clarity, we define several candidate benchmark strategies drawn from the full set of strategies available to block producers. Note that, for block producer t to refer to the last block in the longest chain, she should choose $k_t = t - 1 - t^*(t)$.

- *Strict inclusion:* The strategy profile in which block producer t builds on the longest chain and includes the undesirable transaction whenever possible (Figure 1a). Formally, $a_t = (t - 1 - t^*(t), 1)$ if $e^{\mathcal{C}}(h_t) = 0$ and $a_t = (t - 1 - t^*(t), 0)$ otherwise.

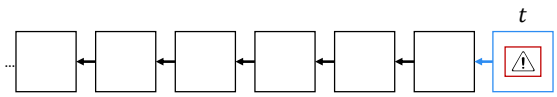
[12] Kahn and Mookherjee (1992) extend the concept of a Coalition-Proof Nash Equilibrium (CPNE) to infinite games. The PCPNE adds the notion of subgame perfection to the CPNE.

[13] A technical subtlety is that these further deviations must themselves be coalition-proof.

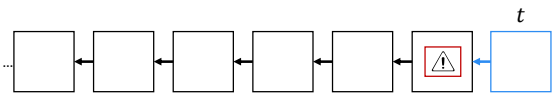
Figure 1: Actions for candidate benchmark strategies

(a) Strict inclusion

Example 1: Include transaction in own block

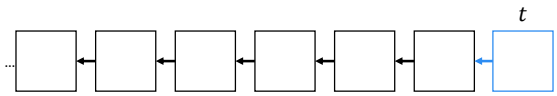


Example 2: Accept block with transaction

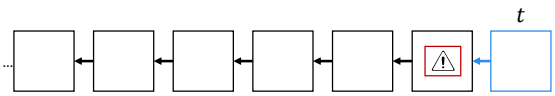


(b) Primary censorship

Example 1: Exclude transaction in own block

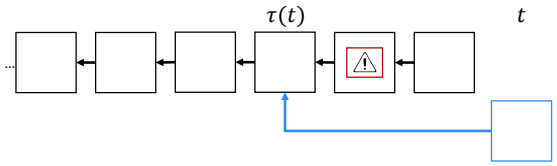


Example 2: Accept block with transaction

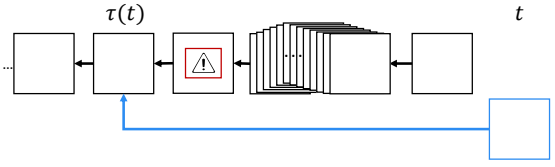


(c) Strict censorship

Example 1: Ignore short chain with transaction

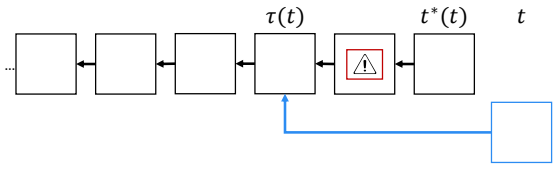


Example 2: Ignore long chain with transaction

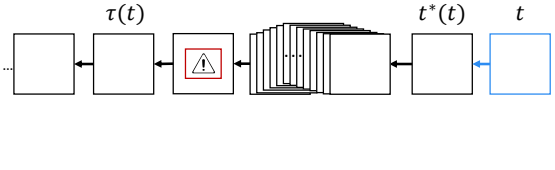


(d) Pragmatic secondary censorship

Example 1: Ignore short chain with transaction



Example 2: Accept long chain with transaction



Note: Slot t reflects the action selected by a block producer who follows the corresponding benchmark strategy. A square with the warning sign represents a block with the undesirable transaction. The blue arrow points to the block to which block producer t attaches her block, ignoring any intermediate blocks.

- *Primary censorship:* The strategy profile in which block producer t builds on the longest chain and neither includes the undesirable transaction nor conducts secondary censorship (Figure 1b). Formally, $a_t = (t-1-t^*(t), 0)$ for any $e^{\mathcal{C}}(h_t)$.

- *Strict censorship:* The strategy profile in which the block producer t does not include the undesirable transaction and links back to the longest chain without the undesirable transaction (Figure 1c). Formally, $a_t = (t-1-\tau(t), 0)$, where $\tau(t)$ is the last block in the longest chain without the undesirable transaction, or, $\tau(t) = \min\left(\arg\max_{s\in[\![0,t-1]\!]}(1-e(h_{s+1}))|T(h_{s+1})|\right)$.

- *Pragmatic secondary censorship:* The strategy profile in which the block producer does not include the undesirable transaction, and conducts secondary censorship only if it could improve her payoff compared to linking back to the longest chain (Figure 1d). Formally, $a_t = (t-1-\tau(t), 0)$ if $\ell > \delta\big(k(h_{\tau(t)+1})+t-1-\tau(t)\big)-\delta(k(h_{t^*(t)+1})+t-1-t^*(t))$, and $a_t = (t-1-t^*(t), 0)$ otherwise.

These candidate benchmark strategies are not necessarily optimal. Strict inclusion does not have to be optimal. Suppose, for instance, that the fee from the transaction is smaller than the loss ($f < \ell$). Then, it cannot be the best response for the current block producer to include the transaction if all past producers excluded the transaction and no future block producer is anticipated to include it. Strict censorship is not necessarily an optimal strategy either. The cost of secondary censorship may be high if the undesirable transaction is "sufficiently deep" in the blockchain in the sense that it would require ignoring a long sequence of blocks (Figure 1c, Example 2), and subsequent block producers may opt to build on the branch that supports inclusion. We will see the optimality of strategies critically depends on the strategies followed by other block producers.

E. *Illustration*

As an illustration, consider the trivial case of a private transaction without the possibility of secondary censorship. With a private transaction, the transaction is only known to the first block producer. Because no other block producer can include the private transaction, this imposes the constraint $e_t = 0$ for all $t \geq 2$. Without the possibility of secondary censorship, block producers also cannot choose to ignore earlier blocks, which imposes the constraint $k_t = 0$ on their action set.

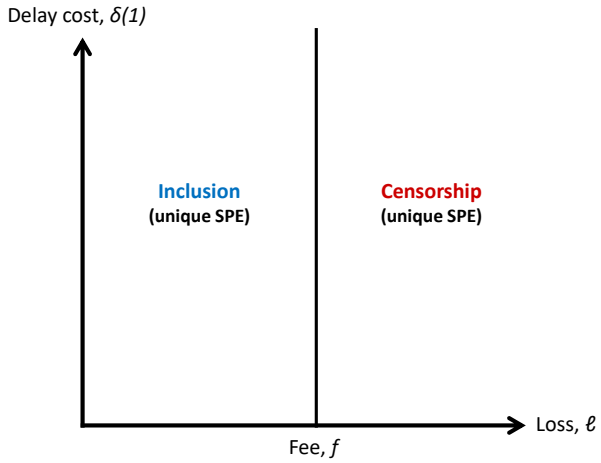
The only choice that remains in the case of a private transaction without secondary censorship is block producer 1's decision on whether to include the undesirable transaction, i.e., $a_1 \in \{(0,0),(0,1)\}$. If she includes the transaction, it will become part of the consensus chain because block producers do not ignore previous blocks. Her choice is effectively between $u_1(0,0) = b$ with censorship and $u_1(0,1) = b + f - \ell$ with inclusion. Other block producers receive a payoff of either $u_{-1} = b$ or $u_{-1} = b - \ell$, where u_{-1} indicates the payoff of all producers other than producer 1. Block producer 1's unique equilibrium play is to include the transaction if $\ell \leq f$ (inclusion) and to exclude it otherwise (censorship). These trivial equilibria are shown in Figure 2a. Regardless of potential delay costs (the vertical axis), inclusion is the unique SPE whenever the loss from inclusion ℓ (the horizontal axis) is smaller than the transaction fee, and censorship otherwise.

V. Analysis Without Secondary Censorship

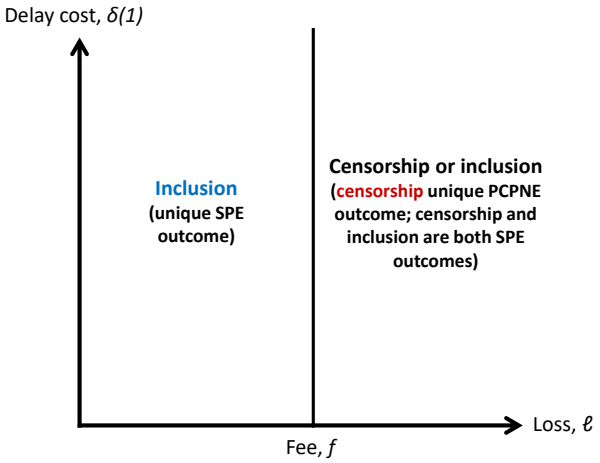
We first characterize the equilibria if block producers do not consider strategic secondary censorship. This corresponds to the standard explanation of the censorship resistance of distributed ledgers that focuses solely on primary censorship. Without secondary censorship, block producers do not ignore previous blocks based on their content, so the action set is restricted to $k_t = 0$ for any t. Block producers may still conduct primary censorship by

Figure 2: Equilibrium Outcomes

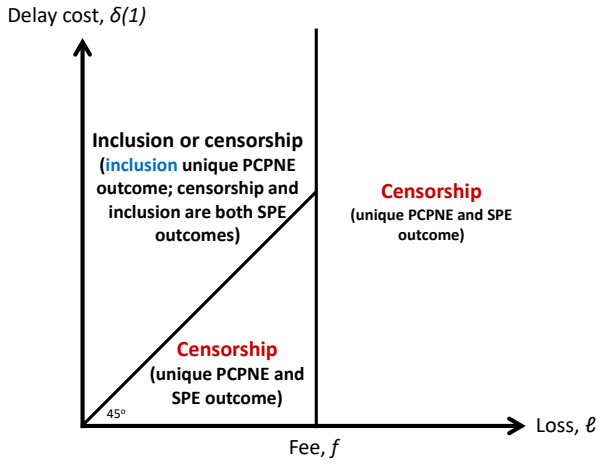
(a) Private transaction without strategic secondary censorship



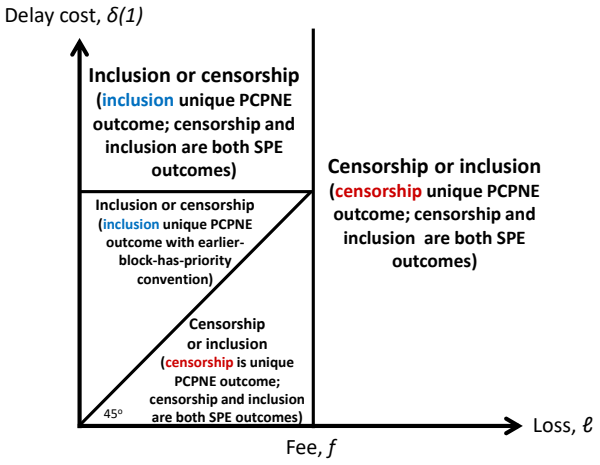
(b) Public transaction without strategic secondary censorship



(c) Private transaction with strategic secondary censorship



(d) Public transaction with strategic secondary censorship



Note: The figures show the different equilibrium outcomes given the parameter values for the transaction fee f, the loss from inclusion of the undesirable transaction in the consensus chain ℓ, and the cost of a failure to reach consensus promptly resulting in an empty slot $\delta(1)$. Strategic secondary censorship refers to the possibility that block producers may strategically choose to ignore blocks that contain an undesirable transaction or blocks that confirm such a block. A "private transaction" is a transaction that cannot be included by subsequent block producers. A "public transaction" can be included by any block producer.

deciding whether or not to include the undesirable transaction in their own block if the chain on which they build does not already include the undesirable transaction.

We will see that, depending on the fee level and the severity of the loss from including the undesirable transaction, the model admits either a unique inclusion equilibrium or multiple equilibria featuring inclusion or censorship. The impact of moving from a private transaction to one that can be included by any block producer is apparent from a comparison of Figures 2a and 2b. The public nature of the transaction introduces inclusion as an additional equilibrium outcome in areas where censorship used to be the unique equilibrium outcome.

We first formalize the result that, without secondary censorship, the undesirable transaction will be included in equilibrium whenever the fee is sufficiently large relative to the negative consequences of inclusion for block producers' payoffs.

Proposition 1 *Consider the case of a public undesirable transaction without strategic secondary censorship. Let $\ell \leq f$. Inclusion is the unique SPE outcome. Block producer 1 includes the undesirable transaction and all other block producers accept her block. Inclusion is also the unique PCPNE outcome.*

The proof is straightforward. All blocks that are proposed must become part of the consensus chain if only primary censorship is considered. Hence, block producers face the choice whether or not to include the undesirable transaction if it has not been included by a previous block producer, but not whether or not to ignore previous blocks. They also do not face the risk that others will exclude their block if they include the undesirable transaction. The acceptance of blocks is guaranteed without secondary censorship.

Because the acceptance of blocks is guaranteed without secondary censorship, a block producer can increase her payoff by including the undesirable transaction in her block if the loss from including the transaction is not greater than the incremental fee, that is, if $\ell \leq f$. Each block producer has incentives to include the transaction if the transaction has not been previously included, regardless of actions of subsequent block producers. The equilibrium

play for each block producer corresponds to the strict inclusion strategy. The outcome of the equilibrium play is that the undesirable transaction will be included in the first block. No block producer has a profitable deviation, so that everyone playing the strict inclusion strategy constitutes an SPE. Also note that no SPE with a censorship outcome can exist, because any block producer could deviate and increase her payoff by switching to the strict inclusion strategy.

The unique inclusion SPE is also perfectly coalition-proof. The first block producer who gets to include the transaction earns the best possible payoff in the equilibrium. Hence, she could never profit by becoming a member of a deviating coalition. The other block producers cannot profit by becoming members of a deviating coalition, because their choice set is constrained such that they have no choice but to accept the block that includes the undesirable transaction.

When the loss from including the undesirable transaction exceeds the transaction fee, then every block producer would be better off if the transaction were censored. However, both the censorship and the inclusion outcome may occur in equilibrium. This is formalized in the following proposition.

Proposition 2 *Consider the case of a public transaction without strategic secondary censorship. Let $\ell > f$. Both inclusion and censorship are possible SPE outcomes. For example, (i) the strategy profile in which all block producers play the strict inclusion strategy is an SPE; (ii) the strategy profile in which all block producers play the primary censorship strategy is an SPE. Censorship is the unique PCPNE outcome.*

It may seem paradoxical that the strategy profile in which every block producer plays the strict inclusion strategy can be an SPE outcome if even the block producer who includes the transaction is worse off than under censorship (since $\ell > f$ implies $b + f - \ell < b$). The apparent paradox arises because the current block producer cannot unilaterally prevent inclusion by censoring the transaction when all the other block producers follow a strict

inclusion strategy. In the inclusion equilibrium, the current block producer anticipates that the next block producer will include the undesirable transaction if she does not include it herself. Given the strict inclusion play of subsequent block producers, she will be better off if she immediately includes the transaction in her own block, thereby earning at least the transaction fee. The same holds true for all subsequent block producers, which is why the profile in which all block producers play a strict inclusion strategy constitutes an SPE.[14]

When the loss from including the undesirable transaction exceeds the transaction fee, then censorship is also a possible SPE outcome. The strategy profile in which all block producers play a primary censorship strategy constitutes an SPE. All block producers earn a payoff b in this equilibrium, which is the maximum possible payoff if $\ell > f$. No block producer has incentives to deviate because the payoff of the block producer who includes the transaction will be reduced from b to $b + f - \ell$.

The censorship equilibrium in which all block producers play a primary censorship strategy is also perfectly coalition-proof. Note that every block producer earns the highest possible payoff b in the censorship equilibrium. Hence, no deviation by a coalition of block producers can exist that increases the payoffs of all its members. Moreover, an equilibrium with an inclusion outcome cannot be perfectly coalition-proof. Consider a coalition consisting of the block producer who includes the transaction and all subsequent block producers. All members of this coalition would be better off by switching to the censorship equilibrium ($b > b+f-\ell$ and $b > b-\ell$). This is a permitted deviation since it is perfectly coalition-proof.

The analysis without secondary censorship can be summarized in three points. First, with public transactions, there always exists an equilibrium (SPE) in which all transactions

[14] As noted above, our analysis focuses on pure equilibria. However, it is worth mentioning there is also a delayed-inclusion SPE in which all block producers play a mixed strategy and include the transaction with a non-zero probability. The mixed strategy is an equilibrium if the transaction is included by each block producer with an independent probability $p = 1-(f/\ell)^{1/q}$. This probability balances the cost of the forgone fee with the benefit of not suffering the loss if the next q block producers do not include the transaction. (The expression for the probability can be obtained from solving $b + f - \ell = b - (1 - (1 - p)^q)\ell$ for p, where the right-hand side reflects the payoff of including the transaction, and the left-hand-side reflects the payoff of not including the transaction, provided that the q next block producers include the transaction with probability p.)

are included, regardless of the losses those transactions impose on block producers. The existence of such an equilibrium seems consistent with the observation that, in practice, many transactions enter the consensus chain, including those that could impose losses on block producers. Second, without strategic secondary censorship, censorship can be an equilibrium outcome only if the loss imposed on a single block producer exceeds the transaction fee earned by that block producer. Hence, in a world where secondary censorship is not possible, individuals could always avoid the censorship of their transaction by setting a transaction fee that is sufficiently high. Thanks to decentralization, this transaction fee only needs to take into account the loss to the block producer who includes the transaction, not the aggregate loss to all block producers. Third, the potential cost of failures to reach consensus measured by the $\delta(\cdot)$-function does not play a role if strategic secondary censorship is not considered.

VI. Analysis with Strategic Secondary Censorship

The possibility of strategic secondary censorship expands the possible action set considerably by allowing block producers to choose any $k_t \in [\![0, t-1]\!]$. Strategic considerations with secondary censorship can be quite complex. For expositional reasons, we guide the reader through the intuition of the relevant strategic considerations by analyzing equilibrium outcomes with strategic secondary censorship using a two-step approach. We first consider equilibrium outcomes with strategic secondary censorship for a private transaction where only producer 1 can include the undesirable transaction.[15] This expands the initially trivial case in Figure 2a with the possibility of secondary censorship. The equilibrium outcomes for this case are shown in Figure 2c. We then consider the case with strategic secondary censorship where any block producer can include the transaction. The equilibrium outcomes for this case are summarized in Figure 2d.

[15]Beyond the expositional point, a private transaction could refer in practice to an MEV transaction that can take advantage of an opportunity only if it is included in the current block (e.g., Capponi et al., 2025; Lehar and Parlour, 2026). There are also parallels with a transaction that is provided to the block producer through a private channel.

A. *Private Transactions and Strategic Secondary Censorship*

The impact of strategic secondary censorship for the private transaction can be well-observed from comparing Figures 2a and 2c. Two main changes appear on the left side of the graph. First, the area where censorship is the unique SPE outcome has been expanded by an additional triangle. Second, inclusion is no longer the unique SPE outcome in the remaining area, although it is the unique PCPNE outcome. Strategic secondary censorship generates multiple equilibria in scenarios in which the loss from including transactions is sufficiently small. These results are formalized in the following proposition.

Proposition 3 *Consider the case of a private undesirable transaction with strategic secondary censorship. Censorship is the unique SPE outcome if and only if at least one of the inequalities* $\ell > f$ *or* $\ell > \delta(1)$ *holds. Censorship is also the unique PCPNE outcome in this scenario. Both inclusion and censorship are possible SPE outcomes if both inequalities* $\ell \leq f$ *and* $\ell \leq \delta(1)$ *hold. Inclusion is the unique PCPNE outcome in this scenario.*

Proof. See Appendix B.2. ■

The additional triangle with censorship as the unique SPE outcome in Figure 2c comes from anticipated secondary censorship by subsequent block producers. The anticipated secondary censorship induces the first block producer to exclude the undesirable transaction from her block. This threat of secondary censorship by subsequent block producers is credible if the delay cost from ignoring a block, $\delta(1)$, is sufficiently small. To see this, consider the situation in which the first block producer chooses to include the undesirable transaction. The next block producer can accept the block or reject it. Assuming that the next block producer's block will be accepted, this involves a trade-off between incurring the loss from including the undesirable transaction ℓ and the cost of an empty slot $\delta(1)$. If $\ell > \delta(1)$, then block producer 2 would prefer to ignore the block of producer 1, earning $b - \delta(1)$ instead of accepting producer 1's block, earning $b - \ell$. If producer 2 were to ignore the block of producer 1, then each subsequent block producer must choose between appending her block to either

the branch including producer 1's block or the branch including producer 2's block. Their payoffs will be higher if they accept the block of producer 2 rather than that of producer 1. Either choice will result in a delay cost of $\delta(1)$ but the block of producer 2 will allow them to avoid the loss ℓ. In other words, the threat of subsequent censorship if block producer 1 were to include the undesirable transaction is credible. Censorship is therefore the unique SPE outcome in this additional triangle.

Inclusion is an equilibrium outcome if both the delay cost and the transaction fee are sufficiently high. Even though secondary censorship would be possible, subsequent block producers will be better off collectively confirming producer 1's block containing the undesirable transaction if the delay cost is sufficiently high. A sufficiently high transaction fee ensures that inclusion will be attractive to producer 1. The inclusion equilibrium requires both $\ell \leq \delta(1)$ and $\ell \leq f$ (upper-left area in Figure 2c). Inclusion is also a perfectly coalition-proof outcome. It provides the highest possible payoff to block producer 1, and, conditional on block producer 1's choice, it provides the highest possible payoff to all other block producers.

The possibility of strategic secondary censorship introduces multiplicity of equilibria in this area where inclusion was the unique equilibrium outcome in the absence of strategic secondary censorship. Even though all block producers are collectively better off confirming a block with the undesirable transaction in this area where the delay cost are relatively high, they individually only have incentives to do so if it does not negatively affect the inclusion of their own block in the consensus chain. In the censorship equilibrium, all subsequent block producers ignore blocks that confirm a chain with the undesirable transaction. Anticipating this, none of the block producers is willing to build upon a chain with the undesirable transaction, which then also induces block producer 1 not to include the transaction. Although this constitutes an SPE outcome, it is not perfectly coalition-proof. The equilibrium requires all block producers to ignore the first block producer if she were to include the undesirable transaction, while, once this happens, the members of a coalition consisting of

all subsequent block producers would each benefit if the coalition collectively switched to the inclusion equilibrium ($\ell \leq \delta(1)$).

These results on strategic secondary censorship can be summarized in two points. First, the threat of strategic secondary censorship may induce primary censorship by block producers. This censorship outcome may be particularly pervasive if the cost of failures to reach consensus is small compared to the loss imposed by the undesirable transaction. In this situation, the censorship outcome will be perfectly coalition-proof. Second, with strategic secondary censorship, it is no longer guaranteed that an undesirable transaction will be included if the transaction fee is set sufficiently high. The block producer who includes the transaction will not be able to benefit from the higher transaction fee if her block is ignored by subsequent block producers, so that a high transaction fee cannot entice her to include the undesirable transaction if she anticipates secondary censorship.

B. Public Transactions and Strategic Secondary Censorship

The present subsection considers the case with both public transactions and strategic secondary censorship. The analysis thus far has considered these aspects separately. We found that, in the absence of secondary censorship, the public nature of transactions can support inclusion equilibria in situations where censorship would otherwise be the unique equilibrium outcome (moving from Figure 2a to Figure 2b). Moreover, if transactions were private, strategic secondary censorship can support censorship equilibria in situations where inclusion would otherwise be the unique equilibrium outcome (moving from Figure 2a to Figure 2c). Secondary censorship also resulted in an additional triangle in Figure 2c in which censorship is the perfectly coalition-proof outcome. The analysis in the present subsection shows that having both public transactions and strategic secondary censorship essentially combines these effects (summarized in Figure 2d).

Before formalizing the results for public transactions and strategic secondary censorship, we introduce the *refined inclusion strategy*, which is identical to the strict inclusion strategy,

except in the situation where the block producer is confronted with multiple longest chains with the same length, but where some of the longest chains exclude the undesirable transaction. The strict inclusion strategy prescribes that the producer simply builds on the oldest branch. The refined inclusion strategy prescribes the producer to build on the oldest branch that excludes the undesirable transaction, and to include the undesirable transaction herself. This will yield her the transaction fee if her block becomes part of the consensus chain.

The following proposition formalizes the results for cases in which censorship is the perfectly coalition-proof outcome.

Proposition 4 *Consider the case of a public transaction with strategic secondary censorship. Suppose at least one of the inequalities $\ell > f$ or $\ell > \delta(1)$ holds. Censorship and inclusion are both SPE outcomes. Censorship is the unique PCPNE outcome. For example, (i) the strategy profile in which all block producers play the refined inclusion strategy is an SPE; and (ii) the strategy profile in which all block producers apply the strict censorship strategy is an SPE and a PCPNE.*

Proof. See Appendix B.3. ■

If the transaction fee is sufficiently low or the cost of delays is sufficiently low, or both, then censorship is an SPE outcome and also a perfectly coalition-proof outcome. This censorship result is similar to the case of a private transaction with strategic secondary censorship. Unlike in the private transaction case, the undesirable transaction can now be included by any block producer due to its public nature. The public nature of the transaction will also support an inclusion equilibrium. If block producer 1 anticipates that all subsequent producers will include the transaction, then she cannot avoid the loss ℓ. She would prefer to include the undesirable transaction herself and collect the incremental fee f, even if it is small. Subsequent block producers have similar incentives to confirm her block. They prefer to avoid the delay cost from censoring the previous block if the loss from inclusion will be incurred anyway. The inclusion outcome is not perfectly coalition-proof. If a coalition consisting of all block producers except the first one were to deviate to a strategy with both

primary and secondary censorship, then a censorship equilibrium would be the result. The censorship outcome leaves all block producers better off if the transaction fee is low (the right side of Figure 2d), and it leaves all block producers except the first one better off if the transaction fee is high (the censorship triangle in Figure 2d). In the latter case, the threat of secondary censorship induces the first block producer to exercise primary censorship.

The following proposition formalizes the results for cases in which inclusion is the perfectly coalition-proof outcome.

Proposition 5 *Consider the case of a public transaction with strategic secondary censorship. Suppose both inequalities $\ell \leq f$ and $\ell \leq \delta(1)$ hold. Censorship and inclusion are both SPE outcomes. Inclusion is the unique PCPNE outcome. For example, (i) the strategy profile where all block producers play the refined inclusion strategy is an SPE and a PCPNE; and (ii) the strategy profile where all block producers play a strict censorship strategy is an SPE.*

Proof. See Appendix B.4. ■

When both the transaction fee and the delay cost are sufficiently high, then inclusion is the perfectly coalition-proof outcome if strategic secondary censorship is considered for a public transaction. This was already the case for private transactions. For public transactions, there is an additional strategic consideration for the first block producer in an inclusion equilibrium. Her incentive to include the transaction arises not only from her own comparison of the transaction fee f to the loss ℓ, but also from the prospect that others will include the undesirable transaction if she does not.

Figure 2d separates the area with the perfectly coalition-proof inclusion outcome into two areas. This separation indicates the different considerations that govern the inclusion outcomes in these areas. The upper rectangle represents the area where, in an inclusion equilibrium, the delay cost is sufficiently high compared to the transaction fee to directly incentivize the second block producer to confirm the block with the undesirable transaction ($\delta(1) > f$). In the upper triangle, more complex considerations arise because the transaction fee exceeds the delay cost. With a public transaction and secondary censorship, the second

block producer may be incentivized to compete for the transaction fee by ignoring the block from the previous block builder. A next block producer can then choose to compete for the transaction fee too, or to build upon any of the earlier blocks. If the next block producer were to confirm one of the two blocks with the undesirable transaction, the choice of which branch to confirm would not directly affect the level of her payoff. The inclusion equilibrium in this scenario relies on the convention that she prefers to build on the earlier branch given equal payoffs. With this convention, there are no incentives for the second block producer to try to compete for the transaction fee because she knows her block will not become part of the consensus chain if she does so.[16]

In summary, with public transactions and secondary censorship, both inclusion and censorship can be equilibrium outcomes. Inclusion is perfectly coalition-proof if the loss from the transaction is low compared to both the transaction fee and the cost of delays in blockchain settlement. Censorship is perfectly coalition-proof if the loss from the undesirable transaction is relatively high.

VII. Effects of Penalties

Jurisdictions may seek to restrict certain on-chain activities by imposing penalties on block producers who include undesirable transactions. The effectiveness of such an approach may be constrained by the limited ability to hold block producers globally accountable. This section evaluates whether a jurisdiction could effectively prevent the processing of an undesirable transaction on a global blockchain when accounting for practical limitations to its enforcement powers.

To evaluate the impact of sanctions, we modify the model in the following way. A governing jurisdiction (or interjurisdictional bloc) imposes a penalty $p > f$ on block producer

[16]The results in this triangle are sensitive to the convention. For example, if the block producer who builds upon a block with the undesirable transaction were to select a block randomly, then there can exist equilibrium outcomes with gaps and delayed inclusion, in which a sequence of multiple block producers include the transaction while ignoring one another's blocks, after which one of the blocks is selected randomly. Such equilibria can be avoided by relying on the convention to build upon the earliest block.

t if the consensus chain reveals that she included the undesirable transaction in her block. Block producers differ in whether the jurisdiction can collect the penalty. We refer to block producers from whom the penalty can be collected as *domestic* block producers. The block producers from whom the penalty cannot be collected are referred to as *foreign* block producers. Block producer types are known.

The threat of penalties introduces heterogeneity in the payoff functions of block producers. The payoff functions of foreign block producers are unchanged from equation (1), but the payoff functions of domestic block producers change to

$$u_t(k_t, e_t, h_{t+q}) = [b - e_t(p - f)]\,\mathbf{1}(t \in T^{\mathcal{C}}(h_{t+q})) - \delta(k^{\mathcal{C}}(h_{t+q})) - e^{\mathcal{C}}(h_{t+q})\ell, \tag{2}$$

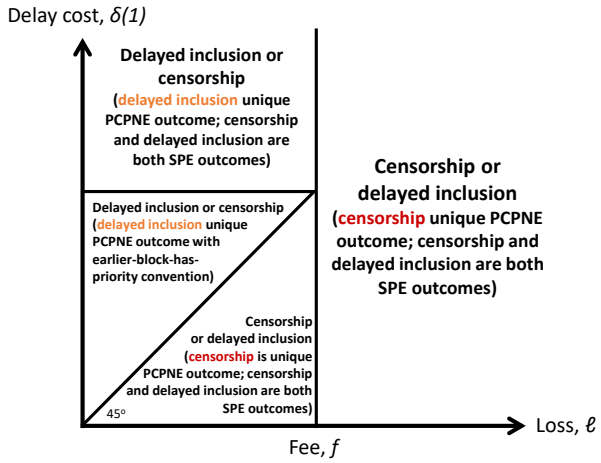
where the penalty p is incurred by the domestic block producer if the consensus chain ultimately reveals that she included the transaction in her block.

We denote the strength of the enforcement powers by $F \geq 1$. Domestic and foreign block producers occur in fixed shares and the types appear in repeated fixed sequences of F slots. The sequence starts with $F-1$ slots that are occupied by domestic block producers and ends with slot F which is occupied by a foreign block producer. The larger F is, the larger the share of block producers over whom the jurisdiction has enforcement powers, which can be calculated as $(F-1)/F$. At its lowest level, $F = 1$, the domestic share equals zero percent. This scenario corresponds to the case in which the jurisdiction has no enforcement powers. The domestic share approaches all block producers as $F \to \infty$. This limit case reflects the situation in which the jurisdiction has global enforcement powers.

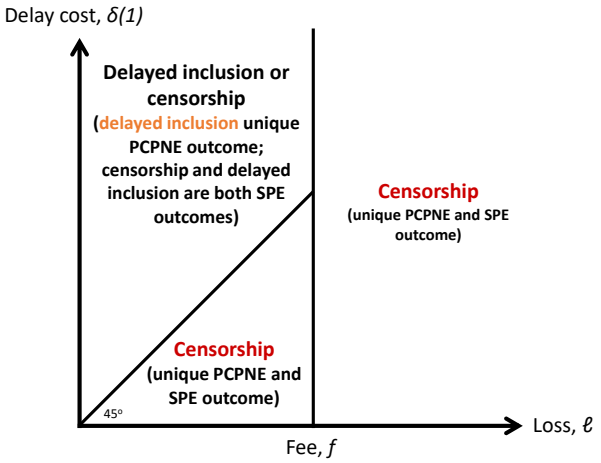
We first provide an explanation of how equilibrium results change depending on the strength of the enforcement powers of the jurisdiction. These results are then summarized in a formal proposition followed by a more general discussion of the effectiveness of penalties in blocking sanctioned transactions.

Figure 3: Equilibrium Outcomes with Sanctions

(a) Limited enforcement powers
$(1 < F < q)$



(b) Strong enforcement powers
$(F \geq q)$



Note: The figures show the different equilibrium outcomes given the enforcement powers of a jurisdiction that imposes fines on block producers who include the undesirable transaction. Parameter F denotes how many blockchain slots it takes until a block producer arrives over whom the jurisdiction has no enforcement powers; q is a large number that reflects the number of slots a block producer must wait before she can exit.

A. No Enforcement Powers $(F = 1)$

Penalties have no impact on the equilibrium results if the jurisdiction that imposes penalties on block producers has no enforcement powers. Without enforcement powers, the payoff functions of block producers are unchanged. The implication is that none of the equilibrium results change. The equilibria will be identical to those in the case of public transactions with strategic secondary censorship in Figure 2d. Inclusion of the undesirable transaction is not prevented. For any set of parameter values, there exists an SPE with an inclusion outcome, and, if both $\ell \leq \delta(1)$ and $\ell \leq f$ hold true, then inclusion will be perfectly coalition-proof.

B. Limited Enforcement Powers $(1 < F < q)$

When penalties are imposed with limited enforcement powers, inclusion of the undesirable transaction may be delayed but inclusion cannot be prevented. The results for limited enforcement powers are summarized in Figure 3a. The only change with respect to the

case in which there are no enforcement powers is that inclusion equilibria are replaced by equilibria with delayed inclusion outcomes. The length of the delay in the delayed inclusion equilibria increases with the strength of the enforcement powers.

Domestic producers have no private incentives to include the transaction as the penalty outweighs the incremental fee revenue. Even when they anticipate that the transaction will be included by a foreign block producer, none of the domestic producers will include the undesirable transaction in their own blocks. Hence, if the government has limited enforcement powers, the penalties must delay the potential inclusion of the undesirable transaction. The stronger the enforcement powers, i.e., the higher the level of F, the longer the expected delay will be before the undesirable transaction will be included in a delayed inclusion equilibrium.

Delayed inclusion equilibria cannot be prevented with penalties if enforcement powers are limited. The incentives to include the undesirable transaction for a foreign block producer are not only affected by her own actions, but also by the anticipated actions of other foreign block producers. If enforcement powers are such that $1 < F < q$, then another foreign block producer could include the undesirable transaction before the payoff to the current block producer materializes. The anticipation of inclusion by the next foreign block producer may induce the current foreign block producer to include the transaction, even if it leaves all foreign block producers worse off, including the current producer. Hence, an SPE exists with the delayed inclusion outcome for any set of parameter values if penalties are imposed with limited enforcement powers.

C. Strong Enforcement Powers ($F \geq q$)

With strong enforcement powers, penalties may prevent inclusion of the undesirable transaction, but only if the undesirable transaction imposes a loss on all block producers that is sufficiently large. Otherwise, delayed inclusion is not prevented. The results for strong enforcement powers are summarized in Figure 3b. The equilibrium outcomes with strong enforcement powers are comparable to the earlier results for a private transaction and

strategic secondary censorship, except that all inclusion equilibria are replaced by delayed inclusion equilibria. The censorship outcome is the only equilibrium outcome if the loss that the undesirable transaction imposes is large, while the delayed inclusion outcome remains perfectly coalition-proof if the loss from the undesirable transaction is small.

The important conceptual difference between limited and strong enforcement powers arises from whether a foreign block producer has to anticipate the actions of other foreign block producers. With enforcement powers such that $F \geq q$, the number of foreign block producers is so low that the payoff of the current foreign block producer will materialize before the next foreign block producer could include the undesirable transaction. None of the domestic block producers will consider including the transaction, but they may still consider imposing secondary censorship. Therefore, a foreign block producer's decision whether to include the undesirable transaction will be comparable to the decision in the private transaction case in which strategic secondary censorship is possible.

D. Formal Sanction Result

The following formal proposition summarizes the results for the different levels of enforcement powers.

Proposition 6 *Consider the case in which a jurisdiction imposes a penalty $p > f$ on domestic block producers who include the undesirable transaction. Let a delayed inclusion outcome be defined as the outcome in which the undesirable transaction is not included until the first foreign block producer appears, after which it is accepted in the consensus chain. Equilibrium outcomes are as follows.*

1. *Suppose $F = 1$. Equilibrium outcomes are identical to the scenario with a public transaction and strategic secondary censorship (Propositions 4 and 5).*

2. *Suppose* $1 < F < q$. *Equilibrium outcomes are identical to the scenario with a public transaction and strategic secondary censorship (Propositions 4 and 5) except that equilibria with inclusion are replaced by equilibria with delayed inclusion.*

3. *Suppose* $F \geq q$. *Equilibrium outcomes are identical to the scenario with a private transaction and strategic secondary censorship (Proposition 3), except that equilibria with inclusion are replaced by equilibria with delayed inclusion.*

Proof. See Appendix B.5. ■

E. Discussion on Effectiveness

Bringing all results together, there are essentially two conditions required for penalties to be effective in eliminating inclusion equilibria. The first requirement is that the jurisdiction must have enforcement powers over an overwhelming share of block producers. The second requirement is that the undesirable transaction must have a sufficiently large negative impact on all block producers, so that cooperating with the jurisdiction is in the foreign block producers' own interests. If either condition is not satisfied, then sanctions simply transform inclusion equilibria into delayed inclusion equilibria. These equilibria are characterized by processing delays for sanctioned transactions, a phenomenon that appears to have occurred following the OFAC designation of selected Ethereum transactions (Brownworth et al., 2024; Wahrstätter et al., 2024).

For any single jurisdiction, it may be difficult to exert enforcement powers over a large share of block producers if the network features substantial geographical diversification (a lower value of F). The degree of international mobility of block producers reduces the extent to which the observed geographical distribution can be treated as fixed. A joint effort of multiple jurisdictions to impose sanctions could be a countervailing force to geographical diversification, thereby bringing a larger fraction of block producers within the enforcement perimeter.

Another factor that increases the difficulty of exerting sufficient enforcement power is a blockchain design that imposes a longer delay between the slot of the current block producer and the moment when she receives her block reward (a higher level of q). A longer delay increases the likelihood that another foreign block producer will include the undesirable transaction. This, in essence, intensifies the "competition" for fees from sanctioned transactions among foreign block producers.

The requirement that the inclusion of sanctioned transactions must impose a sufficiently large loss on all block producers may also be difficult to satisfy. Without such a loss, it is impossible to eliminate inclusion equilibria (if enforcement powers are strong) or even ensure that censorship equilibria are perfectly coalition-proof (if enforcement powers are limited). The threat of fully sanctioning the native token on trading venues or restricting the integration of a given blockchain into payment systems and other financial applications if certain transactions are processed could contribute to such a loss (see also Section II). Such sanctions may undermine the widespread use of a blockchain and its integration in the financial system, yet their effect may be minor when issued by a small jurisdiction or if they target a fringe blockchain.

Even though penalties may not be effective in terms of eliminating delayed inclusion equilibria, one might still argue that formal sanctions by a major jurisdiction could be effective as a device to coordinate on a censorship equilibrium, even if the delayed inclusion equilibrium is perfectly coalition-proof equilibrium. Although this is not impossible, it is important to realize that, if delayed inclusion is perfectly coalition-proof, then it is not possible for a subset of block producers to profitably deviate. In other words, coordinating on a censorship equilibrium would go against the incentives of block producers if the two aforementioned requirements are not satisfied.

VIII. Modifications

A. *Encouraging Inclusion with Side-Payments*

Our analysis does not consider the scenario in which one block producer pays subsequent block producers to induce them to accept the block with the undesirable transaction.[17] Without such side-payments, censorship is the unique perfectly coalition-proof outcome whenever $\ell > \delta(1)$, that is, whenever the loss from including the undesirable transaction exceeds the cost of a single empty slot following secondary censorship. In practice, the cost of a single empty slot may be small, so that secondary censorship may be perfectly coalition-proof whenever the loss from inclusion is non-trivial. The cost of secondary censorship would be higher if multiple subsequent block producers agreed to accept the block with the undesirable transaction in exchange for a side-payment. We analyse the economic feasibility of such "inclusion bonuses" in this subsection.[18]

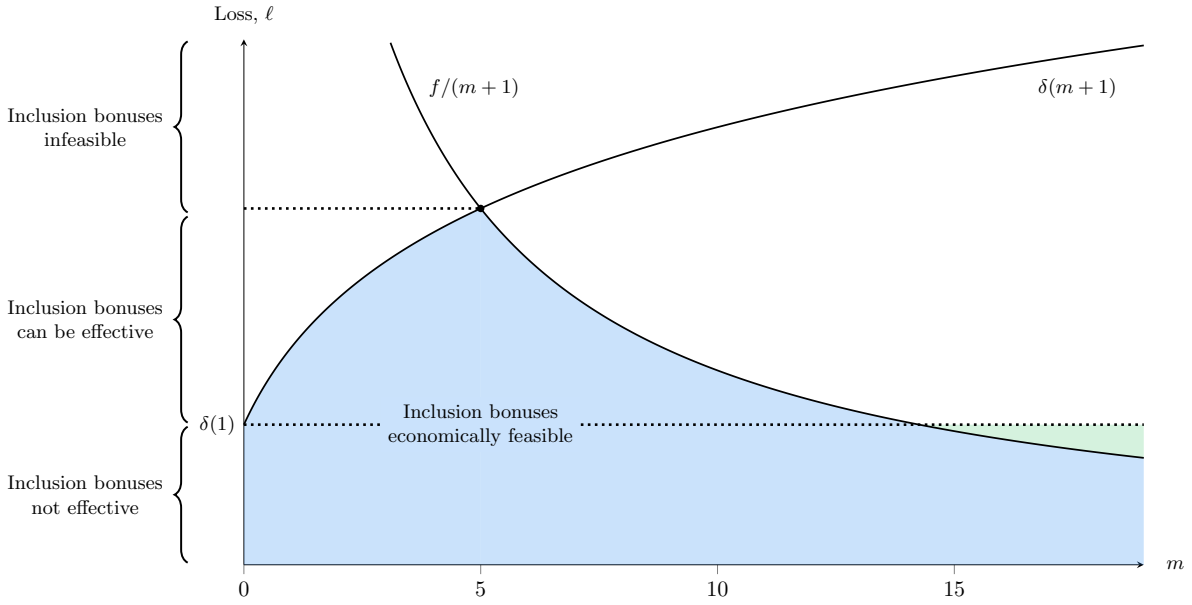
Two conditions need to hold in order for an arrangement with inclusion bonuses to be economically feasible. The first condition is that the transaction fee f must be sufficiently high to cover the losses incurred by both the block producer who includes the transaction and by those who agree to subsequently confirm the block in exchange for a side-payment. Otherwise, it cannot be profitable for a block producer to include the transaction and make sufficiently large side-payments. With m denoting the number of subsequent block producers who receive a side-payment, the first condition requires

$$f \geq (m+1)\ell.$$

[17] Note that our analysis on side-payments is limited in that we do not consider all potential side-payments. For example, we do not consider the scenario in which subsequent block producers promise side-payments to earlier block producers for censoring the undesirable transaction.

[18] Side-payments could be facilitated at the technical level by a smart contract that allows block producers to claim rewards that remain locked into that smart contract until a certain block number is reached.

Figure 4: Encouraging Inclusion with Side-Payments



Note: The figure illustrates the economic feasibility of side-payments to encourage inclusion for the scenario in which $f > \delta(1)$. In the blue-shaded area, it is economically feasible to use side-payments to encourage m subsequent block producers to confirm the block with the undesirable transaction so that all remaining block producers will be worse off in the censorship equilibrium than in the equilibrium that accepts the branch with the undesirable transaction. For $\ell \leq \delta(1)$, which also covers the green area, inclusion bonuses are not effective because inclusion is already the unique PCPNE outcome (Proposition 5). The functional form of $\delta(\cdot)$ is arbitrary (any strictly increasing concave function would do).

The second condition is that ignoring the branch of $m+1$ blocks with the undesirable transaction must be too costly for all the remaining block producers. Otherwise, the censorship equilibrium would still provide them with a better payoff than the inclusion equilibrium. The second condition thus requires

$$\ell \leq \delta(m+1).$$

If there exists a level of $m \geq 0$ for which both conditions hold true, then side-payments can feasibly ensure that, conditional on earlier block producers' actions, the remaining block producers benefit collectively from switching to the inclusion equilibrium.

The blue-shaded area in Figure 4 illustrates parameter combinations for which the two conditions hold. The horizontal axis indicates the number of block producers who receive

a side-payment to confirm the block with the undesirable transaction. The area under the upward-sloping curve identifies the parameter combinations for which the cost from ignoring the branch with the undesirable transaction exceeds the loss from inclusion for the remaining block producers. The area under the downward-sloping curve identifies the parameter combinations for which the transaction fee is sufficiently high to cover the required cost. The figure illustrates the scenario in which $f > \delta(1)$, so that the two curves intersect and inclusion bonuses can be effective. Note that inclusion would already be the unique PCPNE outcome without side-payments in the blue-shaded area below the dotted line, where $\ell \leq \delta(1)$ (Proposition 5).[19] Moreover, given the transaction fee, inclusion bonuses are not feasible if the loss from including the undesirable transaction is too high. Side-payments can be effective for intermediate loss levels because it is economically feasible to make side-payments to a sufficient large sequence of block producers so that the remaining block producers benefit from confirming the branch with the undesirable transaction.

B. Producer of Multiple Consecutive Blocks

Each block producer in our analysis proposes a single block. Various challenges arise if a block producer can propose multiple consecutive blocks. Precise outcomes depend on aspects such as whether block producers know in advance whether they can write multiple consecutive blocks and whether this information is transparent. These aspects depend, among other things, on the consensus protocol. Most proof-of-stake protocols introduce some predictability regarding which block producers will be allowed to produce upcoming blocks, whereas block production occurs in a more random manner for proof-of-work protocols.

To illustrate some of the complexities that arise, we consider a simplified scenario in which a block producer t knows in advance that she is the single block producer who may write

[19] The condition $\ell \leq f$ for Proposition 5 holds true in the blue area because it covers the area where $\ell \leq f/(m+1)$ for $m \geq 0$ and in the green area because the figure covers the case $f > \delta(1)$.

$m+1$ consecutive blocks. We make two observations on how this could promote inclusion outcomes and result in delays due to competition for transaction fees.[20]

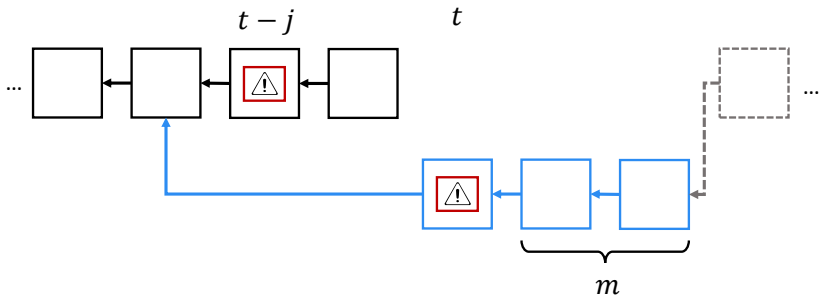
A first observation is that a producer of multiple consecutive blocks may promote inclusion outcomes. In particular, inclusion may be the unique PCPNE outcome for larger levels of the loss ℓ than in our baseline results. A block producer t who knows that she can produce multiple consecutive blocks will internalize the losses for multiple blocks but she can also create a longer branch that includes the transaction. If she were to produce a branch of $m+1$ blocks with the undesirable transaction in the first block of that branch, then an equilibrium with the inclusion outcome would provide all remaining block producers with a better payoff than a censorship equilibrium if $\ell < \delta(m+1)$. Given that the branch with the undesirable transaction will be accepted, it is then optimal for block producer t to produce such a branch if the fee covers the losses, i.e., if $f \geq (m+1)\ell$. These two conditions coincide with the conditions derived in the previous subsection on the economic feasibility of side-payments. Hence, the blue-shaded area in Figure 4 may also be interpreted as the scenarios in which inclusion is the unique PCPNE outcome if producer t knows in advance that she is the single producer who can propose $m+1$ consecutive blocks.[21]

A second observation is that there may be situations where a producer of multiple consecutive blocks could attempt to compete for the transaction fee in an inclusion equilibrium by constructing a competing branch. Suppose there is a single branch that includes the undesirable transaction in block $t-j$. By forking back to block $t-j-1$, multiple-block producer t could ignore the existing branch while adding her own branch of $m+1$ blocks with the undesirable transaction in the first block. In an inclusion equilibrium, it would be

[20] Schär et al. (2026) show that the fee-competition effect may also emerge in a proof-of-work context where the next block producer is unknown. Given a relatively high level of fees in the previous block, it can be optimal for a subsequent block producer to build an alternative branch to compete for those fees if the probability of producing multiple consecutive blocks is sufficiently large for that block producer.

[21] Inclusion is also the unique PCPNE outcome in the green area where $\ell \leq \delta(1)$ but for a different reason. In this area, the transaction fee is less than the loss from inclusion incurred for $m+1$ blocks, so that the multiple-block producer t would be better off in a censorship equilibrium. However, if producer $t+m+1$ responds to censorship by including the transaction, then the remaining block producers would be better off with inclusion. Anticipating inclusion by block producer $t+m+1$, since $f > \delta(1) \geq \ell$, multiple-block producer t includes the transaction immediately to collect the fee.

Figure 5: Competing for the Fee with Multiple Consecutive Blocks



Note: A block producer t who can add $m+1$ consecutive blocks causes a delay in the consensus chain of j empty slots by competing for the transaction fee.

optimal for the remaining block producers to accept the new branch rather than the original branch if the delay cost were smaller, which will be the case for $j < m+1$. The multiple-block producer would be better off by competing for the transaction fee if the transaction fee exceeds the delay cost associated with her branch, which increases in proportion to the number of blocks she can propose, that is, if $f \geq (m+1) \cdot \delta(j)$. Depending on whether the earlier block producer $t-j$ could anticipate losing the competition, she might wish to leave inclusion of the undesirable transaction entirely to the multiple-block producer.

IX. Concluding Remarks

This paper analyzed censorship resistance of blockchains through the lens of a game-theoretic framework. We introduced the concept of secondary censorship as the choice to selectively ignore blocks produced by others if those blocks contain or confirm technically valid but undesirable transactions. The threat of secondary censorship may induce block producers to pre-emptively engage in primary censorship of transactions. A high transaction fee is not sufficient to guarantee the inclusion of a transaction if block producers consider strategic secondary censorship.

The model also reveals that coordination can play an important role in determining censorship outcomes. Equilibria with inclusion and censorship outcomes generally coexist.

Inclusion is perfectly coalition-proof if a transaction imposes no or sufficiently small losses on block producers, whereas censorship is perfectly coalition-proof if the loss imposed on block producers is sufficiently large. Censorship is thus most prone to arise in scenarios in which coordination is practically feasible and substantial financial interests are at stake for block producers. Two examples satisfying both conditions are the intervention following the 2016 DAO attack on the Ethereum blockchain, where a smart contract prevented the adversary from moving diverted funds for several weeks, and the more recent discussions regarding the freezing of quantum-vulnerable Bitcoin addresses in anticipation of future attacks.

References

O. Acuna. To Freeze or not to Freeze: Satoshi and the $440 Billion in Bitcoin Threatened by Quantum Computing. Coindesk Report (February 23), 2026a. URL http://tiny.cc/oui0101.

O. Acuna. Bitcoin Developer Jameson Lopp Says It's Better to Freeze 5.6 Million BTC Than Let Hackers Have Them. Coindesk Report (April 15), 2026b. URL http://tiny.cc/ut42101.

Y. Amoussou-Guenou, B. Biais, M. Potop-Butucaru, and S. Tucci-Piergiovanni. Committee-based Blockchains as Games between Opportunistic Players and Adversaries. *Review of Financial Studies*, 37(2):409–443, 2024.

I. Barmes, B. Bosch, and O. Haalstra. Quantum Computers and the Bitcoin Blockchain. Deloitte Perspective, 2020. URL http://tiny.cc/9xi0101.

B. Berger, E.W. Felten, A. Mamageishvili, and B. Sudakov. *Economic Censorship Games in Fraud Proofs*, pages 670–687. Association for Computing Machinery, New York, NY, 2025.

B.D. Bernheim, B. Peleg, and M.D. Whinston. Coalition-Proof Nash Equilibria I. Concepts. *Journal of Economic Theory*, 42(1):1–12, 1987.

L. Bertucci, H. Jahanshahloo, and S. Scharnowski. Do Cryptocurrency Investors Care About Quantum Risks? *Working Paper*, 2025.

B. Biais, C. Bisière, M. Bouvard, and C. Casamatta. The Blockchain Folk Theorem. *Review of Financial Studies*, 32(5):1662–1715, 2019.

Bitcoin Core Developers. Bitcoin Core: `consensus.h` (line 19). https://github.com/bitcoin/bitcoin/blob/de925455c8025fc1f75d65d981c28b9dfa20e9f7/src/consensus/consensus.h#L19, 2026a. Accessed: 2026-05-25.

Bitcoin Core Developers. Bitcoin Core: `validation.cpp` (line 3271). https://github.com/bitcoin/bitcoin/blob/7dea464d6b51a69bd99a0451be8aaf3a26313eb6/src/validation.cpp#L3271, 2026b. Accessed: 2026-06-28.

A. Brownworth, J. Durfee, M. Lee, and A. Martin. Regulating Decentralized Systems: Evidence from Sanctions on Tornado Cash. *Federal Reserve Bank of New York Staff Report*, 1112, 2024.

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

V. Buterin. CRITICAL UPDATE Re: DAO Vulnerability. *Ethereum Foundation Blog Post (June 16)*, 2016. URL https://blog.ethereum.org/2016/06/17/critical-update-re-dao-vulnerability.

C. Cachin, J. Mićić, N. Steinhauer, and L. Zanolini. Quick Order Fairness. In I. Eyal and J. Garay, editors, *Financial Cryptography and Data Security*, pages 316–333. Springer International Publishing, 2022.

A. Capponi, R. Jia, and K.Y. Wang. Maximal Extractable Value and Allocative Inefficiencies in Public Blockchains. *Journal of Financial Economics*, 172:104132, 2025.

J. Chiu and T.V. Koeppl. The Economics of Cryptocurrencies: Bitcoin and Beyond. *Canadian Journal of Economics*, 55(4):1762–1798, 2022.

C. Detrio. EIP-779: Hardfork Meta: DAO Fork, 2017. URL https://eips.ethereum.org/EIPS/eip-779.

I. Dragos, K. Karantias, and W.J. Knottenbelt. Bitcoin Crypto: Bounties for Quantum Capable Adversaries. In P. Pardalos, I. Kotsireas, Y. Guo, and W. Knottenbelt, editors, *Mathematical Research for Blockchain Economy*, pages 9–25. Springer International Publishing, 2020.

Z. Ebrahimi, M. Guennewig, B. Routledge, and A. Zetlin-Jones. Achieving Consensus on Blockchains. *Working Paper*, 2026.

Ethereum Foundation. Staking Withdrawals. https://launchpad.ethereum.org/en/withdrawals, 2025. Accessed: 2026-05-25.

D. Fudenberg and J. Tirole. *Game Theory*. MIT Press, Cambridge, Massachusetts, 1991.

P. Garimidi, J. Neu, and M. Resnick. Multiple Concurrent Proposers: Why and How. *Working Paper*, 2025.

R.J. Garratt. A Distributed Version of Repugnance as a Constraint on Markets. *Federal Reserve Bank of New York: Liberty Street Economics*, September 2015. URL https://tinyurl.com/4cxnr863.

R.J. Garratt and M.R.C. Van Oordt. Why Fixed Costs Matter for Proof-of-Work–Based Cryptocurrencies. *Management Science*, 69(11):6482–6507, 2023.

O. Godbole. Arthur Hayes Floats the Idea of Rolling Back Ethereum Network to Negate $1.4B Bybit Hack, Drawing Community Ire. Coindesk Report (22 February), 2025. URL http://tiny.cc/7jv5101.

V. Gramoli, Z. Lu, Q. Tang, and P. Zarbafian. AOAB: Optimal and Fair Ordering of Financial Transactions. In *2024 54th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, pages 377–388, 2024.

H. Hałaburda, Z. He, and J. Li. An Economic Model of Consensus on Distributed Ledgers. *NBER Working Paper*, 29515, 2021.

H. Hałaburda, G. Haeringer, J.S. Gans, and N. Gandal. The Microeconomics of Cryptocurrencies. *Journal of Economic Literature*, 60(3):971–1013, 2022.

L. Heimbach, L. Kiffer, C. Ferreira Torres, and R. Wattenhofer. Ethereum's Proposer-Builder Separation: Promises and Realities. In *Proceedings of the 2023 ACM on Internet Measurement Conference*, pages 406–420, New York, NY, USA, 2023. Association for Computing Machinery.

K. John, B. Monnot, P. Mueller, F. Saleh, and C. Schwarz-Schilling. Economics of Ethereum. *Journal of Corporate Finance*, 91:102718, 2025a.

K. John, T.J. Rivera, and F. Saleh. Proof-of-Work versus Proof-of-Stake: A Comparative Economic Analysis. *Review of Financial Studies*, 38(7):1955–2004, 2025b.

C.M. Kahn and D. Mookherjee. The Good, the Bad, and the Ugly: Coalition Proof Equilibrium in Infinite Games. *Games and Economic Behavior*, 4(1):101–121, 1992.

A. Kiayias, E. Koutsoupias, and A.P. Stouka. Incentives Against Power Grabs or How to Engineer the Revolution in a Pooled Proof of Stake System. In *2021 IEEE International Conference on Decentralized Applications and Infrastructures (DAPPS)*, 2021.

A. Lehar and C.A. Parlour. Microstructure of Blockchain Cryptocurrency Markets. *Annual Review of Financial Economics*, 2026. ISSN 1941-1367.

J. Lopp, C. Papathanasiou, I. Smith, J. Ross, S. Vaile, and P.L. Dallaire-Demers. Post Quantum Migration and Legacy Signature Sunset, 2026. URL http://tiny.cc/ew42101.

J.A. Meyer. Second Thoughts on Secondary Sanctions. *University of Pennsylvania Journal of International Law*, 30(3):905–968, 2009.

T.C. Morgan, C. Syropoulos, and Y.V. Yotov. Economic Sanctions: Evolution, Consequences, and Challenges. *Journal of Economic Perspectives*, 37(1):3–29, 2023.

A. Narayanan, J. Bonneau, E. Felten, A. Miller, and S. Goldfeder. *Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction*. Princeton University Press, 2016.

B. Peleg. On Perfectly Coalition-Proof Nash Equilibria. In M. Majumdar, editor, *Equilibrium and Dynamics: Essays in Honour of David Gale*, pages 259–268. Palgrave Macmillan, 1992.

S. Reback and J. Crawley. Binance to Delist Monero Privacy Token; XMR Slides. Coindesk Report (6 February), 2024. URL http://tiny.cc/ak94101.

F. Schär, D. Thürkauf, and D. Yermack. Unruly by Design: Fee Volatility and Strategic Attacks in Bitcoin Mining. *NBER Working Paper*, 35443, 2026.

T. Thiery, F. D'Amato, J. Ma, B. Monnot, T. Tsao, J. Kaufmann, and J. Song. EIP-7805: Fork-Choice Enforced Inclusion Lists (FOCIL), 2024. URL https://eips.ethereum.org/EIPS/eip-7805.

S. Wadhwa, J. Ma, T. Thiery, B. Monnot, L. Zanolini, F. Zhang, and K. Nayak. AUCIL: An Inclusion List Design for Rational Parties. *Working Paper*, 2025.

A. Wahrstätter, J. Ernstberger, A. Yaish, L. Zhou, K. Qin, T. Tsuchiya, S. Steinhorst, D. Svetinovic, N. Christin, M. Barczentewicz, et al. Blockchain Censorship. In *Proceedings of the ACM Web Conference 2024*, pages 1632–1643, 2024.

A. Illustration of a PCPNE

To illustrate the difference between an SPE and a PCPNE, consider the game tree for a finite two-stage game with three players in Figure 6. The figure departs from the standard representation of an extensive-form game to emphasize the simultaneous moves by players 2 and 3 once player 1 has played action L. Note that any Nash equilibrium requires players 2 and 3 to play the same action in response to player 1 choosing L (we consider pure strategies).

The game in Figure 6 has two SPE outcomes and a unique PCPNE. The following strategy profiles produce the two SPE outcomes:

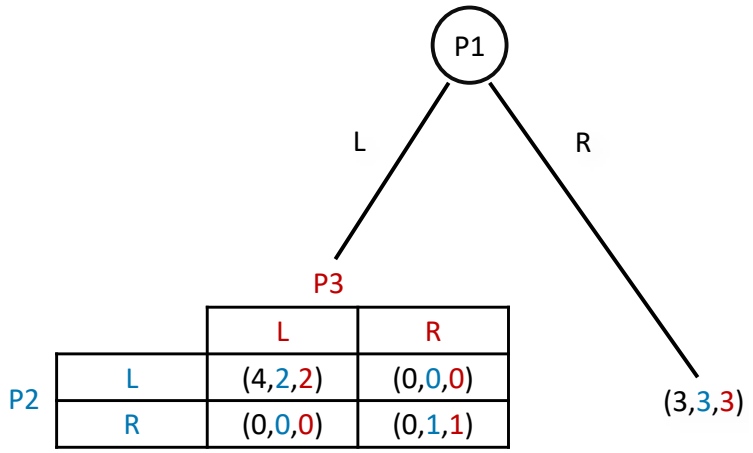
1. Players 2 and 3 both respond by choosing R if player 1 chooses L. Player 1 chooses L if players 2 and 3 both respond by playing L, and R otherwise. Given these strategies, player 1 will choose R. The SPE outcome is the payoff vector $(3, 3, 3)$.

2. Players 2 and 3 both respond by choosing L if player 1 chooses L. Player 1 chooses L if players 2 and 3 both respond by choosing L, and R otherwise. Given these strategies, all players will choose L. The SPE outcome is the payoff vector $(4, 2, 2)$.

It is straightforward to verify that neither strategy profile admits a profitable unilateral deviation.

Strategy profile 1 does not correspond to a PCPNE. The strategies of players 2 and 3 in the second stage are not perfectly coalition-proof. Once player 1 has chosen L, players 2 and 3 would end up with a payoff of 1 if they were to follow through on both playing R. However, once player 1 made her choice L, they would both be better off forming a coalition that deviates to playing L, which would improve their payoffs to 2. This would be a permitted deviation for the coalition, because it is not possible for players 2 and 3 to improve their individual payoffs by deviating further from the coalition's proposed strategy. Strategy profile 1 is not a PCPNE given the existence of this profitable permitted deviation.

Figure 6: Illustration of Difference between SPE and PCPNE



Note: The figure shows the game tree for a two-stage game with three players. Player 1 moves first, after which players 2 and 3 move simultaneously. The game has two SPE outcomes and a unique PCPNE.

Strategy profile 2 does correspond to a PCPNE. To prove this, we consider all possible coalitions. Note that no single-player coalitions can profitably deviate because the strategy profile corresponds to an SPE. Moreover, player 1 cannot be part of a multiplayer coalition whose deviation would yield her a better payoff, because her equilibrium payoff is already the highest possible outcome in the game. This leaves us with deviations by a multiplayer coalition consisting of players 2 and 3. There also is no deviation by which a coalition consisting of players 2 and 3 could profitably deviate because the equilibrium provides them both with the highest possible payoffs conditional upon player 1 choosing L.

The PCPNE refinement in this game eliminates what could be considered as a "non-credible threat" of players 2 and 3 (that is, to both respond with R if player 1 chooses L). The SPE refinement is not able to eliminate the non-credible threat in this game. Not carrying out the threat involves a deviation by multiple players, whereas an SPE only rules out profitable single-player deviations.

B. Proofs

B.1. General

No-unnecessary-empty-slot (NUES) rule:

The number of branches in the tree representing the extensive-form game expands quickly with the number of block producers. For example, the third block producer may choose $k_3 \in \{0, 1, 2\}$; the fourth block producer may choose $k_4 \in \{0, 1, 2, 3\}$, et cetera. To prune back the tree, we first derive the no-unnecessary-empty-slot (NUES) rule, which eliminates branches that represent dominated choices.

Block producers have incentives to avoid situations where proposed blocks are ignored *unnecessarily* because ignored blocks increase the cost $\delta(k^{\mathcal{C}}(h_{t+q}))$. By comparing payoffs in equation (1), we derive the following NUES rule:

Block producer t must choose $k_t = 0$ unless at least one of the following exemptions occurs:

1. choosing $k_t \neq 0$ results in the exclusion of a block with the undesirable transaction (i.e., a change from $e^{\mathcal{C}}(h_{t+q}) = 1$ to $e^{\mathcal{C}}(h_{t+q}) = 0$);

2. choosing $k_t \neq 0$ results in the inclusion of the current producer's block (i.e., a change from $\mathbf{1}(t \in T^{\mathbb{C}}(h_{t+q})) = 0$ to $\mathbf{1}(t \in T^{\mathbb{C}}(h_{t+q})) = 1$); or,

3. choosing $k_t \neq 0$ results in the collection of the fee for the undesirable transaction, which requires the producer's block to be accepted (i.e., a change from $e_t\mathbf{1}(t \in T^{\mathbb{C}}(h_{t+q})) = 0$ to $e_t\mathbf{1}(t \in T^{\mathbb{C}}(h_{t+q})) = 1$).

If none of these exemptions occur, then the payoff from $k_t = 0$ immediately dominates that of $k_t \neq 0$. Only if one of these exemptions applies *may* it be optimal to choose $k_t \neq 0$.

B.2. Proof of Proposition 3

Proposition 3 concerns the case of a private transaction and secondary censorship. In this case, after eliminating all the branches that violate the NUES rule, the remaining choices are as follows. Producer 1 decides whether to include or censor the undesirable private transaction. Producer 1 is the only one who can include the private transaction. If she includes the undesirable transaction, then each subsequent block producer chooses whether to support inclusion or censorship. Support is given by building on the longest existing chain that is consistent with either inclusion or censorship.

Censorship is the unique SPE outcome if $\ell > f$ or $\ell > \delta(1)$:

Consider the decision of producer 1. With a private transaction, no one else can include the transaction. Suppose producer 1 censors the undesirable transaction. Then no one else can include it, and, given the NUES rule, every subsequent producer will simply confirm her block. Hence, $a_1 = (0,0)$ yields a certain payoff $u_1 = b$. Suppose producer 1 includes the undesirable transaction. In her best-case scenario, subsequent block producers accept her block, in which case $a_1 = (0,1)$ would yield a payoff $u_1 = b + f - \ell$. When comparing the payoffs, $a_1 = (0,0)$ dominates the best possible payoff for $a_1 = (0,1)$ if $f < \ell$. Hence, $f < \ell$ is a sufficient condition for primary censorship by producer 1.

Consider the decision of producer 2. In the scenario of a private transaction, $e_2 = 0$, so we only have to consider the decision of producer 2 to ignore blocks. Ignoring a block is permitted only if an exemption to the NUES rule applies. Consider each of the potential exemptions for producer 2:

Exemption 1 may apply if producer 1 included the undesirable transaction, $a_1 = (0,1)$. The question is whether subsequent producers will accept producer 2's block if she ignores the block of producer 1, so that it results in a change from $e^{\mathcal{C}}(h_{t+q}) = 1$ to $e^{\mathcal{C}}(h_{t+q}) = 0$. The answer is affirmative. Consider the decision by block producer 3. If producer 2's block ignores producer 1's block, then their blocks are mutually exclusive. Producer 3 cannot avoid

an empty slot in the blockchain. Producer 3 could accept producer 2's block, but this would mean ignoring producer 1's block; alternatively, ignoring producer 2's block also implies an empty slot. The NUES rule ensures that producer 3 would not ignore more than one block if the two preceding blocks are mutually exclusive while one block supports censorship. Hence, producer 3 effectively has a choice between a payoff $u_3 = b - \delta(1)$ (if block 2 becomes part of the consensus chain and $e^{\mathcal{C}}(h_{t+q}) = 0$), or a payoff of $u_3 = b - \delta(1) - \ell$ (if block 1 becomes part of the consensus chain and $e^{\mathcal{C}}(h_{t+q}) = 1$). Since $\ell > 0$, producer 3 would accept producer 2's block if that block ignores the block with the undesirable transaction of producer 1. Similar decisions hold for all subsequent block producers, but with larger costs of delays if they opt to link back to producer 1's block. Hence, producer 2's choice to ignore producer 1's block plays a pivotal role in blocking the undesirable transaction from the consensus chain if producer 1 chooses $a_1 = (0, 1)$. Given the pivotal role of producer 2, producer 2's decision to ignore block 1 with the undesirable transaction will depend on comparing the cost of an empty slot with the loss following inclusion. Producer 2's best response to inclusion, i.e., $a_1 = (0, 1)$, is secondary censorship, i.e., $a_2 = (1, 0)$, if $\ell > \delta(1)$.

Exemption 2 may apply if producer 2's acceptance of producer 1's block with the undesirable transaction may induce subsequent producers to ignore both blocks. Consider the decision by producer 3 if $a_1 = (0, 1)$ and $a_2 = (0, 0)$. Given the NUES rule, producer 3 will either support censorship or inclusion based on the longest possible chain. Producer 3 must choose between ignoring 2 blocks (support censorship) and ignoring no blocks (support inclusion). Assume her block would be accepted by subsequent producers if she ignores two blocks. Then producer 3 would have incentives to censor earlier blocks if $\ell > \delta(2)$. If the condition $\ell > \delta(2)$ holds, then $\ell > \delta(1)$ holds true, since $\delta(2) > \delta(1)$. Hence, if this exemption could apply, then producer 2 would already have ignored the block of producer 1 under exemption 1. Hence, exemption 2 does not result in additional scenarios in which producer 1's block will be ignored.

Exemption 3 cannot apply for producer 2 in case of a private transaction because this case imposes the constraint $e_2 = 0$.

It follows from the relevant exemptions that, if $a_1 = (0, 1)$ and $\ell > \delta(1)$, producer 2's best response is secondary censorship: $a_2 = (1, 0)$. Subsequent producers will confirm her block. Hence, if $\ell > \delta(1)$, then producer 1's block will be ignored if it includes the undesirable transaction (credible threat of secondary censorship), resulting in a payoff $u_1 = -\delta(1)$. Hence, it is optimal for producer 1 to impose primary censorship under the credible threat of secondary censorship, $a_1 = (0, 0)$, which results in a payoff $u_1 = b$. We have also established that producer 1 is incentivized to conduct primary censorship if $\ell > f$, which would be confirmed by subsequent producers. Hence, censorship is the unique SPE outcome if $\ell > f$ or $\ell > \delta(1)$. The payoff to block producers in the censorship outcome equals $u_t = b, \forall t$.

Perfectly Coalition-Proof. The censorship outcome is also a perfectly coalition-proof outcome. We show this by proving that, for any of the stages, there are no deviations of coalitions where all members of the coalition profit.

Consider stage 1. Suppose there exists a coalition of *multiple* block producers that could deviate such that inclusion would be the outcome. Then this coalition must necessarily include at least one block producer $t > 1$. The payoffs of such a coalition member would equal $u_t = b - \ell$ when the intended deviation results in an inclusion outcome, which is lower than the payoff $u_t = b$ in a censorship outcome. The only coalition where all members could possibly profit is the single-player coalition consisting of producer 1. Since the censorship outcome is an SPE, a deviation by a single player cannot possibly be optimal (one-stage deviation principle). Hence, there are no deviations by coalitions that benefit all members for stage 1.

For any subsequent stages, the argument follows from the observation that the censorship outcome achieves the best possible payoff for all remaining block producers. Hence, for any stage $T > 1$, conditional upon the play by block producers $t < T$, no deviations exist for

coalitions consisting of block producers $t \geq T$, that could possibly improve the payoffs of the members of that coalition.

Multiple SPE outcomes if both $\ell \leq f$ and $\ell \leq \delta(1)$

1. SPE with inclusion outcome:

Consider the potential inclusion equilibrium in which producers play the following strategies: producer 1 includes the undesirable transaction; and, all subsequent producers refer back to the last block in the chain that includes the undesirable transaction, or simply to the last block if no such chain exists. The payoffs if everyone follows these strategies are $u_1 = b + f - \ell$ and, for all $t > 1$, $u_t = b - \ell$.

We show this is an SPE using the one-stage deviation principle. The only deviation available to producer 1 is to exclude the undesirable transaction. Given the strategies of all other players, this yields a payoff $u_1 = b$, which is less than the equilibrium payoff since $\ell \leq f$. Given the NUES rule, the only deviation available to any producer $t > 1$ is to link back to the initial state before producer 1's block, i.e., $a_t = (t - 1, 0)$. Given the equilibrium play by subsequent block producers, who refer back to the last block in the chain that includes the undesirable transaction, this will result in block t being excluded from the consensus chain, resulting in a payoff of $u_t = -\ell - \delta(1)$. This is strictly less than the equilibrium payoff $u_t = b - \ell$. Hence, no individual block producer has a profitable deviation.

Perfectly coalition-proof: The described inclusion equilibrium is also perfectly coalition-proof.

Consider a potential coalition that includes producer 1. This coalition cannot improve the payoffs of all players. In the equilibrium, producer 1 receives the payoff $u_1 = b + f - \ell$, which is the highest possible payoff since $\ell \leq f$.

Consider any coalition excluding producer 1. Such a coalition cannot improve the payoffs of all players once producer 1 includes the undesirable transaction, $a_1 = (0, 1)$. Equilibrium play yields the best possible payoff $u_t = b - \ell$ to all subsequent producers conditional upon $a_1 = (0, 1)$. The best possible deviation is a deviation by a coalition consisting of all block

producers $t > 1$ who decide to censor the first block while accepting one another's blocks. This deviation would yield them $u_t = b - \delta(1)$. This is lower than the equilibrium payoff $u_t = b - \ell$ since $\ell \leq \delta(1)$. At any stage $t > 2$, similar deviations would earn even lower payoffs due to the larger number of blocks that need to be ignored to censor block 1.

Since no permitted deviations exist at any stage, the inclusion equilibrium is *perfectly* coalition-proof.

2. SPE with censorship outcome:

Consider the potential censorship equilibrium in which block producers play the following strategies: producer 1 excludes the undesirable transaction; and, all subsequent producers refer back to the last block in the chain that excludes the undesirable transaction. If everyone adheres to these strategies, the payoffs are $u_t = b$ for any t.

We show these strategies constitute an SPE using the one-stage deviation principle. The only deviation available to producer 1 is to include the undesirable transaction. Given the strategies of all other block producers, subsequent block producers will ignore her block, resulting in a payoff of $u_1 = -\delta(1)$. Given the NUES rule, no deviations are available to any producer $t > 1$ if producer 1 does not include the undesirable *private* transaction. Hence, not a single block producer has a profitable deviation.

Not perfectly coalition-proof: Any equilibrium with a censorship outcome cannot be perfectly coalition-proof. To see this, consider the following deviation at stage 2 if producer 1 were to include the undesirable transaction. In the censorship equilibrium, all subsequent block producers $t > 1$ would censor block 1 while taking into account the NUES rule, yielding them a payoff $u_t = b - \delta(1)$. Instead, a profitable deviation for a coalition consisting of all producers $t > 1$ would be to accept producer 1's block, yielding all of them a payoff $u_t = b - \ell$. The deviation towards inclusion is profitable since $\ell \leq \delta(1)$. The deviation is perfectly coalition-proof because it provides all remaining producers with the best possible payoff conditional upon $a_1 = (0, 1)$. When anticipating this profitable deviation, it is optimal

for producer 1 to play $a_1 = (0, 1)$ yielding a payoff $u_1 = b + f - \ell$ rather than $a_1 = (0, 0)$ yielding a payoff $u_1 = b$, since $\ell \leq f$.

B.3. Proof of Proposition 4

Thanks to the NUES rule, some actions are immediately dominated. We restrict the presentation to the actions that are not dominated. Producer 1 can decide whether to include or censor the undesirable transaction. Unlike in the previous proposition, if producer 1 does not include the undesirable transaction, subsequent producers can still decide to include it due to the public nature of the transaction. If a producer decides to include an undesirable transaction, subsequent block producers either accept the inclusion or they censor the block and perform primary censorship themselves.

Multiple SPE outcomes if $\ell > f$ or $\ell > \delta(1)$

The proposition states that, contrary to the case of a private undesirable transaction, inclusion can be an SPE outcome with a public transaction and secondary censorship. This SPE outcome coexists with the censorship SPE outcome.

1. SPE with censorship outcome:

Consider the potential censorship equilibrium in which block producers play the following strategy: all producers exclude the undesirable transaction and refer back to the last block in the chain that excludes the undesirable transaction. If everyone follows this strategy, the payoffs are $u_t = b$ for any block producer t. Deviations to ignore blocks without the undesirable transaction are dominated under the NUES rule. We still need to analyze deviations where a block producer includes the undesirable transaction.

We show that collectively playing the censorship strategy constitutes an SPE using the one-stage deviation principle. The only deviation available to a producer t that is not covered by the NUES rule is to include the undesirable transaction. Given the strategies of all other block producers, subsequent block producers will ignore her block, resulting in a payoff of

$u_t = -\delta(1)$ for the block producer who deviates by including the transaction, while her equilibrium payoff is $u_t = b$. This holds true for any single block producer. Hence, no single-deviation for any block producer exists that improves her payoff.

Perfectly coalition-proof: The censorship equilibrium is also a perfectly coalition-proof outcome if at least one of the inequalities $\ell > f$ or $\ell > \delta(1)$ holds. We show this by proving that, for any of the stages, there are no coalition deviations in which all members of the coalition profit.

Consider potential deviations where the undesirable transaction is included at stage 1.

First, consider all potential coalitions that include block producer 1, whose collective deviation would result in an inclusion outcome. The only manner in which producer 1 could possibly profit from a deviation would be if she included the undesirable transaction. Any coalition of *multiple* block producers must necessarily include at least one block producer $t > 1$. The payoff of such a coalition member would equal $u_t = b - \ell$ when a deviation results in an inclusion outcome, which is lower than the payoff $u_t = b$ in the censorship outcome. Hence, the only coalition where all members could possibly profit is the single-player coalition consisting of producer 1. Since the censorship outcome is an SPE, a deviation by a single player cannot possibly be optimal. Hence, there are no deviations of coalitions where all members profit if block producer 1 is a member of the coalition.

Second, consider any potential coalition that excludes block producer 1. At stage 1, they could deviate in their response to block producer 1. Equilibrium play requires them to ignore block producer 1 if she includes the undesirable transaction. The only available deviation is to accept block 1 in response to block producer 1 including the undesirable transaction. If such a deviation were successful in achieving the inclusion outcome, then the deviation would yield each member of the coalition $b - \ell$, which is lower than the equilibrium payoff $b - \delta(1)$ when they ignore a block that includes the undesirable transaction, since $\ell > \delta(1)$. Hence, at stage 1, there are no deviations of any coalitions excluding block producer 1 where all members of the coalition profit.

For any subsequent stages, the argument follows from repeating the above argument by considering deviations by coalitions that include or exclude block producer 2, 3, et cetera.

2. SPE with inclusion outcome:

Consider the candidate inclusion equilibrium in which all block producers play the refined inclusion strategy. In this situation, the producers play according to the following rules:

- Producer 1 includes the undesirable transaction.
- Producer 2 refers back to the block of producer 1, and includes the undesirable transaction if it has not been included by block producer 1.
- Any producer $t > 2$ applies the following rules.
  - If there is a single branch, then refer back to the last block of that branch and include the undesirable transaction if it is not included in the branch.
  - If there are multiple branches, then refer back to the branch with the smallest number of missing blocks and include the undesirable transaction if it has not been included in that branch.
  - If there are multiple branches with the same number of missing blocks, then refer back to the last block in the branch that excludes the undesirable transaction and include the undesirable transaction.
  - If there are ties between multiple branches based on these rules, then apply the earlier-block-has-priority convention to select a branch.

The equilibrium payoffs if all block producers follow these strategies are $u_1 = b + f - \ell$ for producer 1 and $u_t = b - \ell$ for any block producer $t > 1$. We prove these strategies constitute an SPE using the one-stage deviation principle.

Consider any producer $t > 1$. Given the NUES rule, there are two potential deviations for any block producer $t > 1$:

The *first potential deviation* is to ignore the chain with the undesirable transaction *without* including the transaction in her own block, effectively imposing secondary censorship by linking back to the initial state before producer 1's block, i.e., $a_t = (t-1, 0)$ ("censoring the undesirable transaction"). In the inclusion equilibrium, this deviation is dominated by the equilibrium payoff $b - \ell$: Block producer t's payoff from the deviation is at most the larger of $u_t = -\ell - \delta(1)$, which is the payoff if her deviating block is ignored, and $u_t = b - \ell - \delta(t-1)$, which is the payoff if her deviating block (without the transaction) becomes part of the consensus chain. Regardless of whether her block is accepted, the actions of subsequent block producers will still result in an inclusion outcome.

Whether a deviating block $t > 1$ that censors the undesirable transaction would be accepted depends on the value of t. For any producer $t > 2$, the strategies of the subsequent block producers would result in her deviating block being ignored, leading to a payoff $-\ell - \delta(1)$. For block producer $t = 2$, the strategies of the subsequent block producers would result in the acceptance of her deviating block. The strategy of block producer 3 would be to accept the block of block producer 2 and to include the undesirable transaction in block 3. This branch will then be accepted by all subsequent producers because it would be the branch with the least missing blocks. The deviation would yield block producer 2 a payoff of $b - \ell - \delta(1)$, less than the payoff $b - \ell$ from equilibrium play.

The *second potential deviation* is to include the undesirable transaction while ignoring the existing branch with the undesirable transaction, i.e., to play $a_t = (t-1, 1)$ ("competing for the undesirable transaction"). Regardless of whether her block is accepted, the actions of subsequent block producers will still result in an inclusion outcome. Moreover, we will see that, given the strategies of the other block producers, this deviation is dominated too.

Block producer t's payoff from this deviation could be at most either $u_t = -\ell - \delta(1)$, which is the worse payoff if her deviating block will be ignored, or $u_t = b + f - \ell - \delta(t-1)$, which is the payoff if her deviating block will become part of the consensus chain. Conditional upon the block producer t's deviating block becoming part of the consensus chain, the

deviation could result in a payoff higher than the equilibrium payoff if $f > \delta(t-1)$, but not if $f < \delta(1)$. Hence, if the fee is sufficiently small such that $f < \delta(1)$, then the deviation must be dominated by the equilibrium payoff $b - \ell$ for any block producer $t > 1$, regardless of whether their deviating block would be accepted. This scenario concerns the area north of the point where the diagonal intersects with the vertical line in Figure 2d. Equilibrium play must dominate this deviation in this area.

Alternatively, consider the scenario in which the transaction fee is sufficiently large such that $f \geq \delta(1)$. This scenario concerns the area south of the point where the diagonal intersects with the vertical line in Figure 2d. The deviation would then be profitable for block producer 2 if she were sure her block would become part of the consensus chain. Her problem is that it will not. Suppose block producer 2 deviates by including the transaction. Block producer 3 could then choose between linking back to the branch consisting of block 1, or the branch consisting of block 2, or she could also join the competition for the transaction fee by creating another branch. Note that if she links back to any branch, then the earlier-block-has-priority convention induces her to prefer the branch of block producer 1 rather than that of a subsequent block producer. The same would apply to a block producer 4 if block producer 3 were instead to create her own branch. Hence, given the earlier-block-has-priority convention, deviating by creating a competing branch that includes the undesirable transaction next to an existing branch must must ultimately cause block producers to ignore the deviating block, leaving the deviating producer with a payoff $-\delta(1) - \ell$. Hence, no block producer has incentives to compete for the transaction fee once it is included by a previous block producer given the earlier-block-has-priority convention.

Let us now turn to producer 1. Her only possible deviation would be to censor the undesirable transaction. The payoff from this deviation, given the equilibrium play of the other block producers who will include the transaction, would be $b - \ell$. This is less than the equilibrium payoff $b + f - \ell$. Hence, no individual block producer has a profitable deviation. The candidate inclusion equilibrium is an SPE.

Not perfectly coalition-proof: Inclusion equilibria are not perfectly coalition-proof if at least one of the inequalities $\ell > f$ or $\ell > \delta(1)$ holds. For either inequality, we show this by contradiction.

First, consider the case in which $\ell > f$. Suppose there were a perfectly coalition-proof equilibrium with inclusion as an outcome. In such an equilibrium, the payoff of any producer t who includes the transaction can be at most $u_t = b + f - \ell < b$, while other producers would have a payoff of at most $u_{t'} = b - \ell < b, \forall t' \neq t$.

Consider the deviation by the coalition consisting of all block producers where they now play strategy profiles according to the SPE with the censorship outcome. The SPE with the censorship outcome is itself perfectly coalition-proof and hence a permitted deviation. This is a profitable deviation, because, given the deviation, the payoff of all block producers equals b, which exceeds each block producer's payoff in the inclusion equilibrium. This is a contradiction.

Second, consider the case in which $\ell > \delta(1)$. Suppose there were a perfectly coalition-proof equilibrium in which the undesirable transaction is included in block t. The payoff of the producers would be at most $u_t = b + f - \ell$ and $u_{t'} = b - \ell, \forall t' \neq t$.

Consider the deviation by a coalition consisting of all block producers after t. If each member of this coalition deviates to strict censorship, then each deviating producer would get a payoff of $b - \delta(1)$ instead of $b - \ell$ if t still plays inclusion. Since $\ell > \delta(1)$, each member of the deviating coalition would be better off. Moreover, this deviation is permitted because all the deviating producers will receive the maximal possible payoff conditional on player t's inclusion of the undesirable transaction. Player t's best response to the deviating coalition is to switch to also censoring the undesirable transaction, resulting in a payoff of $b > -\delta(1)$ and a censorship outcome. This contradicts the claim that the equilibrium in which the transaction is included at time t can be perfectly coalition-proof if $\ell > \delta(1)$. The argument holds for any t.

If both $\ell > f$ and $\ell > \delta(1)$ hold true, either argument can be followed. In summary, there cannot exist a perfectly coalition-proof equilibrium with the inclusion outcome if at least one of the inequalities $\ell > f$ or $\ell > \delta(1)$ holds.

B.4. Proof of Proposition 5

The proposition concerns the case in which both $\ell \leq f$ and $\ell \leq \delta(1)$ hold true. In Figure 2d, this case is reflected by the areas covered by the upper-left rectangle and the upper triangle. Both inclusion and censorship are SPE outcomes in these areas, whereas only an equilibrium with an inclusion outcome can be perfectly coalition-proof. As for the previous proposition, we restrict the proof to actions that are not immediately dominated according to the NUES rule.

Multiple SPE outcomes if both $\ell \leq f$ and $\ell \leq \delta(1)$

1. SPE with inclusion outcome:

The proof is a repetition of the proof of the SPE with an inclusion outcome for Proposition 4 in the previous subsection.

Perfectly coalition-proof: The inclusion equilibrium is perfectly coalition-proof.

Consider the situation in which producer 1 includes the undesirable transaction, $a_1 = (0, 1)$. Then, no profitable deviations can exist at any stage $T > 1$ because equilibrium play yields the best possible payoff $u_t = b - \ell$ to all subsequent producers conditional upon $a_1 = (0, 1)$. For example, at stage 2, the relevant deviation permitted under the NUES rule is by a coalition of all block producers $t > 1$ who decide to censor the first block and not include the undesirable transaction while accepting one another's blocks. This deviation would yield them $u_t = b - \delta(1)$. This is lower than the payoff under equilibrium play $u_t = b - \ell$ since $\ell \leq \delta(1)$. At any stage $t > 2$, similar deviations would earn even lower payoffs due to the larger number of blocks that need to be ignored to censor block 1.

Now, consider deviations at stage 1. At stage 1, a coalition of all block producers $t > 1$ could consider deviating by "threatening" to censor the block from producer 1 if she includes

the transaction. If producer 1 were to censor the transaction because of this threat, it would improve the payoffs of all coalition members from $u_t = b-\ell$ to $u_t = b$. However, this deviation by the coalition does not constitute a credible threat: Once producer 1 plays $a_1 = (0, 1)$, it would be optimal for the coalition consisting of all block producers $t > 1$ to return to accepting the block with the undesirable transaction as shown above. Hence, no permitted deviations exist at any stage, and the described equilibrium with the inclusion outcome must be perfectly coalition-proof.

Finally, consider the situation in which producer 1 excludes the undesirable transaction, $a_1 = (0, 0)$. Following the equilibrium play, block producer 2 will include the undesirable transaction and all remaining block producers will accept 2's block. In this case, the payoff of block producer 2 is $b + f - \ell$ and the payoff of block producers $t > 2$ is $b - \ell$. At stage 3, the only relevant deviation, thanks to the NUES rule, will be by a coalition of all block producers $t > 2$ who decide to censor the block of the second block producer and to not include the undesirable transaction while accepting one another's blocks. This deviation would yield them at best $u_t = b-\delta(1)$ (by accepting block producer 1's block). This is lower than the payoff they receive if they accept block producer 2's block $(b - \ell)$ since $\ell \leq \delta(1)$. The argument holds similarly for any stage $t > 2$ if the first $t-1$ block producers play $(0, 0)$ and block producer t plays $(0, 1)$. Inclusion is then guaranteed, and no coalition can improve its payoff once inclusion by producer t occurs.

2. SPE with censorship outcome:

Consider the potential censorship equilibrium in which block producers play the following strategies: producer 1 excludes the undesirable transaction; and, all subsequent producers refer back to the last block in the chain that excludes the undesirable transaction. If everyone adheres to these strategies, the payoffs are $u_t = b$ for any t.

We show this constitutes an SPE using the one-stage deviation principle. Given the NUES rule, the only deviation available to any producer t is to include the undesirable transaction. Given the strategies of all other block producers to ignore branches that include

the transaction, her block will be ignored, resulting in a payoff of $u_t = -\delta(1)$. This is lower than the payoff from equilibrium play $u_t = b$. Hence, any block producer t prefers to censor the undesirable transaction because no profitable one-stage deviation exists.

Not perfectly coalition-proof: Censorship equilibria are not perfectly coalition-proof. To see this, consider the following deviation at stage 2 if producer 1 were to include the undesirable transaction. In a censorship equilibrium, all subsequent block producers $t > 1$ would have to censor block 1 while taking into account the NUES rule, yielding them a payoff of $u_t = b - \delta(1)$. Instead, a profitable deviation for a coalition consisting of all producers $t > 1$ would be to accept producer 1's block, yielding a payoff of $u_t = b - \ell, \forall t > 1$. The deviation towards inclusion is profitable due to the high delay cost such that $\ell \leq \delta(1)$. The deviation is perfectly coalition-proof because it provides all remaining producers with the best possible payoff conditional upon $a_1 = (0, 1)$. Anticipating such a deviation by the coalition, it is optimal for producer 1 to include the undesirable transaction and play $a_1 = (0, 1)$ yielding a payoff of $u_1 = b + f - \ell$ rather than $a_1 = (0, 0)$ yielding a payoff $u_1 = b$, since $\ell \leq f$.

B.5. Proof of Proposition 6

The sequence of $F - 1$ domestic block producers has no incentives to include the undesirable transaction. Regardless of whether the undesirable transaction becomes part of the consensus chain, they are better off if they do not include it themselves. If a domestic block producer were to include the undesirable transaction and inclusion were accepted, then her payoff would be $b + f - p - \ell$. Since $f < p$, this is a lower payoff than her payoff if the undesirable transaction is not included in the consensus chain, b, or if another block producer were to include the transaction in the consensus chain, $b - \ell$ (a subsequent block producer has no incentives to ignore a block produced by someone who conducts primary censorship).

Foreign block producers may have incentives to include the undesirable transaction. Their payoffs will depend on whether subsequent block producers respond by accepting inclusion or by conducting strategic secondary censorship. Moreover, foreign block producers have to

account for the actions by subsequent foreign producers who can also decide to include the undesirable transaction or not. We restrict the presentation to actions that are not ruled out by the NUES rule.

Case 1 (no enforcement powers, $F = 1$):

There are no domestic producers if $F = 1$. None of the payoff functions in the model change. The existing equilibria and proofs are identical to those for the model without penalties (Propositions 4 and 5).

Case 2 (limited enforcement powers, $1 < F < q$):

Multiple SPE outcomes if $\ell > f$ or $\ell > \delta(1)$

The proposition states that censorship and delayed inclusion are both SPE outcomes if at least one of the inequalities $\ell > f$ or $\ell > \delta(1)$ holds, while censorship is the unique PCPNE outcome.

1. SPE with censorship outcome:

The proof of the existence of an SPE with the censorship outcome is a replication of the proof of the existence of an SPE with the censorship outcome for Proposition 4.

Perfectly coalition-proof: The censorship equilibrium is a perfectly coalition-proof outcome if at least one of the inequalities $\ell > f$ or $\ell > \delta(1)$ holds. We show this by proving that, for any of the stages, there are no deviations of coalitions where all members of the coalition profit.

Consider potential deviations where the undesirable transaction is included at stage F.

The first $F - 1$ producers are domestic. Their equilibrium payoffs are b, which is their highest possible payoff. None of these block producers can improve their payoffs by becoming part of a coalition. We therefore move to the first foreign block producer.

Consider all potential coalitions that include block producer F, whose collective deviation would result in an inclusion outcome. The only manner in which producer F could possibly profit from a deviation would be if she included the undesirable transaction. Any coalition

of *multiple* block producers must necessarily include at least one block producer $t > F$ who accepts the inclusion of producer F. The payoffs of any such coalition member would be at most $u_t = b - \ell$ when a deviation results in an inclusion outcome, which is lower than the payoff $u_t = b$ in the censorship outcome. Hence, the only coalition where all members could possibly profit is the single-player coalition consisting of producer F. Since the censorship outcome is an SPE, a deviation by a single player cannot possibly be optimal. Hence, there are no deviations of coalitions where all members profit if block producer F is a member of the coalition.

Finally, consider any potential coalition that excludes block producer F. At stage F, block producers could deviate in their response to block producer F. Equilibrium play requires them to ignore block producer F if she includes the undesirable transaction. The only available deviation that could result in inclusion is to accept block F in response to block producer F including the undesirable transaction. If such a deviation were successful in achieving the inclusion outcome, then the deviation would yield each member of the coalition $b - \ell$. If $\ell > \delta(1)$, the payoff is lower than the equilibrium payoff $b - \delta(1)$ where they ignore the block that includes the undesirable transaction. Alternatively, if $\ell \leq \delta(1)$ and $\ell > f$, then the payoff would be higher if they accepted the block with the undesirable transaction, but then block producer F would have no incentives to include the transaction given the response by the coalition. Hence, at stage F, there are no profitable deviations that result in an inclusion outcome for any coalitions excluding block producer F.

For any subsequent stages, the argument follows from repeating the above argument by considering deviations by coalitions that include or exclude block producer $2F$, $3F$, et cetera.

2. SPE with delayed inclusion outcome:

Consider the candidate inclusion equilibrium in which domestic block producers play the primary censorship strategy, and all foreign block producers play the refined inclusion strategy. In this situation, the producers play according to the following rules:

- Domestic producers do not include the undesirable transaction.

- Any producer t applies the following rules.
  - If there is a single branch, then refer back to the last block of that branch and, if the producer is a foreign producer, include the undesirable transaction if it is not included in that branch.
  - If there are multiple branches, then build on the last block of the branch with the smallest number of missing blocks and, if she is a foreign producer, include the undesirable transaction if it has not been included in that branch.
  - If there are multiple branches with the same number of missing blocks, then refer back to the last block in the branch that excludes the undesirable transaction and, if she is a foreign producer, include the undesirable transaction.
  - If there are ties between multiple branches based on these rules, then apply the earlier-block-has-priority convention to select a branch.

If all block producers follow these strategies, then the undesirable transaction will be included by block producer F and accepted by subsequent block producers. The equilibrium payoffs are $u_F = b + f - \ell$ for producer F and $u_t = b - \ell$ for any block producer $t \neq F$. We prove these strategies constitute an SPE using the one-stage deviation principle.

Consider any producer $t > F$. Given the NUES rule, there are two relevant potential deviations for a block producer $t > F$. The first potential deviation is to ignore the chain with the undesirable transaction *without* including the transaction in her own block. The second potential deviation is to include the undesirable transaction while ignoring the existing branch with the undesirable transaction (this deviation is trivially dominated for domestic producers).

Consider the *first potential deviation.* Producer t ignores the chain with the undesirable transaction *without* including the transaction in her own block, effectively imposing secondary censorship by linking back to the initial state before producer F's block, i.e., $a_t = (t - F, 0)$ ("censoring the undesirable transaction"). In the delayed-inclusion equi-

librium, this deviation is dominated by the equilibrium payoff $u_t = b - \ell$. Regardless of whether her block is accepted, the actions of subsequent block producers will still result in an inclusion outcome by time $t + q$. The next foreign producer will include the undesirable transaction, which happens before $t + q$ since $F < q$, and her block will be accepted by the other block producers. Hence, block producer t's payoff from the deviation is at most the larger of $u_t = -\ell - \delta(1)$, which is the payoff if her deviating block is ignored, or $u_t = b - \ell - \delta(t - F)$, which is the payoff if her deviating block becomes part of the consensus chain. Both are less than the equilibrium payoff $u_t = b - \ell$.

Consider the *second potential deviation*. It involves including the undesirable transaction while ignoring the existing branch with the undesirable transaction, i.e., to play $a_t = (t-F, 1)$ ("competing for the undesirable transaction"). This deviation is only relevant for foreign producers. Given the strategies of the other block producers, this deviation is dominated too. Block producer t's payoff from the deviation is at most the larger of $u_t = -\ell - \delta(1)$, which is the payoff if her deviating block is ignored, or $u_t = b + f - \ell - \delta(t - F)$, which is the payoff if her deviating block becomes part of the consensus chain. Regardless of whether her block is accepted, the actions of subsequent block producers will still result in an inclusion outcome.

Conditional upon foreign block producer t's deviating block becoming part of the consensus chain, the deviation could result in a higher payoff than the equilibrium payoff only if $f > \delta(t - F)$. If $f < \delta(1)$, then the fee is sufficiently small such that the deviation must be dominated by the equilibrium payoff $b - \ell$ for any foreign block producer $t > F$, regardless of whether her deviating block would be accepted. Alternatively, consider the scenario in which the transaction fee is sufficiently large such that $f \geq \delta(t - F)$. The deviation would then be profitable for block producer t if she were sure her block would become part of the consensus chain. Her problem is that it will not. Suppose, for example, that foreign block producer $2F$ deviates by including the transaction. Block producer $2F + 1$ could then choose between linking back to the branch containing block F, resulting in a delay cost $\delta(1)$, or the branch

consisting of block $2F$, resulting in a delay cost $\delta(F)$. Since $\delta(1) < \delta(F)$ for $F > 1$, it would be the best response for subsequent block producers to link to the branch containing block F. If she links back to any branch, then the earlier-block-has-priority convention induces her to prefer referring back to the branch of block producer F rather than that of a subsequent block producer. The same reasoning applies to foreign block producers $3F$, $4F$, and so forth. Hence, no foreign block producer has incentives to compete for the transaction fee once it is included by a previous foreign block producer.

Not perfectly coalition-proof: Equilibria with delayed inclusion are not perfectly coalition-proof if at least one of the inequalities $\ell > f$ or $\ell > \delta(1)$ holds. The proof is a replication of the proof that inclusion is not a PCPNE outcome in Proposition 4.

Multiple SPE outcomes if both $\ell \leq f$ and $\ell \leq \delta(1)$

The proposition states that censorship and delayed inclusion are both SPE outcomes if both $\ell \leq f$ and $\ell \leq \delta(1)$, while delayed inclusion is the unique PCPNE outcome.

1. SPE with censorship outcome:

The proof of the existence of an SPE with the censorship outcome is a replication of the proof of the existence of an SPE with the censorship outcome for Proposition 5.

Not perfectly coalition-proof: Censorship equilibria are not perfectly coalition-proof if both $\ell \leq f$ and $\ell \leq \delta(1)$ hold. To see this, consider the following deviation at stage $F+1$ if producer F were to include the undesirable transaction. In a censorship equilibrium, all subsequent block producers $t > F$ would have to censor block F while taking into account the NUES rule, yielding them a payoff of $u_t = b - \delta(1)$. Instead, a profitable deviation for a coalition consisting of all producers $t > F$ would be to accept producer F's block, yielding a payoff of $u_t = b - \ell, \forall t > F$. The deviation towards delayed inclusion is profitable due to the high delay cost such that $\ell \leq \delta(1)$. The deviation is perfectly coalition-proof because it provides all remaining producers with the best possible payoff conditional upon $a_F = (0, 1)$, and $a_t = (0, 0)$, for all $t < F$. Anticipating such a deviation by the coalition, it is optimal

for producer F to include the undesirable transaction and play $a_F = (0, 1)$, resulting in an inclusion outcome and yielding her a payoff of $u_F = b + f - \ell$, which, since $\ell \leq f$, is more than the payoff $u_F = b$ from playing $a_F = (0, 0)$.

2. SPE with delayed inclusion outcome:

The proof of the existence of an SPE with delayed inclusion is a replication of the proof of the SPE with delayed inclusion outcome above for the case in which at least one of the inequalities $\ell > f$ or $\ell > \delta(1)$ hold.

Perfectly coalition-proof: Delayed inclusion is a PCPNE outcome if both $\ell \leq f$ and $\ell \leq \delta(1)$ hold.

Consider the situation in which producer F includes the undesirable transaction, $a_F = (0, 1)$. Then, no profitable deviations can exist at any stage $t > F$ because equilibrium play yields the best possible payoff $u_t = b - \ell$ to all subsequent producers conditional upon $a_F = (0, 1)$. For example, at stage $F+1$, the relevant deviation to consider due to the NUES rule will be by a coalition of all block producers $t > F$ who decide to censor block F and not to include the undesirable transaction while accepting one another's blocks. This deviation would yield them $u_t = b - \delta(1)$, which is lower than the payoff under equilibrium play since $\ell \leq \delta(1)$. At any stage $t > F + 1$, similar deviations would earn even lower payoffs because censoring block F requires them to ignore a longer branch.

Finally, consider deviations at stage F. At stage F, a coalition of all block producers $t > F$ could consider deviating by "threatening" to censor the block from producer F if she includes the transaction. If producer F were to censor the transaction because of this threat, it would improve the payoffs of all coalition members from $u_t = b - \ell$ to $u_t = b$. However, this deviation by the coalition is not permitted. Once producer F plays $a_F = (0, 1)$, it would be optimal for the coalition consisting of all block producers $t > F$ to deviate return to accepting the block with the undesirable transaction as shown above. In other words, the deviation does not constitute a credible threat.

Finally, consider deviations at stage $t < F$. No producer $t < F$ would be willing to include the undesirable transaction as a member of a coalition that deviates to inclusion since $p > f$. Moreover, if none of the producers $t < F$ includes the transaction, then the NUES rule implies that they form a single branch in which they all build upon the previous block.

Hence, no permitted deviations exist at any stage t, and the described equilibrium with the delayed inclusion outcome must be perfectly coalition-proof.

Case 3 (strong enforcement powers, $F \geq q$):

When the jurisdiction has strong enforcement powers, consecutive foreign block producers are separated by at least $q-1$ domestic block producers. Consider a foreign block producer t. Under the payoff function for foreign producers in equation (1), the actions of any subsequent foreign block producer t' will have no impact on the payoff of foreign block producer t, since $t' \geq t + q$ (as $t' \geq t + F$). All $q - 1$ subsequent block producers are domestic, and none of them has incentives to include the undesirable transaction as $p > f$. Hence, they act as if they do not have access to the undesirable transaction. Any foreign block producer t will therefore act as if no other producer will include the transaction for the next $q - 1$ blocks. Domestic block producers can, however, engage in secondary censorship.

The analysis follows as in the case of a private transaction with strategic secondary censorship (Proposition 3). The first block producer to have an economic incentive to add the transaction is block producer F. Thus, inclusion equilibria become equilibria with *delayed* inclusion. If at least one of the inequalities $\ell > f$ or $\ell > \delta(1)$ holds, censorship is the unique SPE outcome. Censorship is also the unique PCPNE outcome in this scenario. If both inequalities $\ell \leq f$ and $\ell \leq \delta(1)$ hold, delayed inclusion and censorship are both possible SPE outcomes. Delayed inclusion is the unique PCPNE outcome in this scenario. The proof follows the same lines as that of Proposition 3.