

Consensus Change Standards

*A Legal and Technical Framework
for Bitcoin Protocol Governance*

Asaf Fulks

California State Bar No. 343622

Solo Bitcoin Miner · Full Node Operator



a Fulks, Inc. company

Fifth Edition · July 2026

asaffulkslaw.com

CONSENSUS CHANGE STANDARDS

*A Legal and Technical Framework
for Bitcoin Protocol Governance*

Fifth Edition — July 2026

WRITTEN BY

ASAF FULKS, J.D.

California State Bar No. 343622

Admitted, U.S. District Court, Central District of California

asaffulkslaw.com

THE FORUM PRESS

a Fulks, Inc. company | California

theforumpress.com

CONSENSUS CHANGE STANDARDS

A Legal and Technical Framework for Bitcoin Protocol Governance

Copyright © 2026 by Asaf Fulks. All rights reserved subject to the license below.

License: Creative Commons Attribution 4.0 International (CC BY 4.0). You are free to share, copy, redistribute, adapt, remix, transform, and build upon this work for any purpose, including commercial use, provided you give appropriate credit to the author, indicate if changes were made, and do not suggest the author endorses your use.

Full license terms: creativecommons.org/licenses/by/4.0

Fifth Edition, July 2026

Fourth Edition, June 2026

Third Edition, June 2026 (v3.1 revision)

Second Edition, May 2026

First Edition, April 2026 (revised May 2026)

Published by The Forum Press, a Fulks, Inc. company

California | theforumpress.com

Available at: asaffulkslaw.com

Cite as: Asaf Fulks, *Consensus Change Standards: A Legal and Technical Framework for Bitcoin Protocol Governance* (5th ed. 2026), asaffulkslaw.com.

Notice and Disclaimer. This document is provided for informational and educational purposes only. It does not constitute legal advice and is not a solicitation for legal services or attorney advertising. The legal analysis herein draws primarily on California state law and federal authority; readers in other jurisdictions should not assume it applies to their circumstances. Nothing in this document creates an attorney-client relationship between the author and any reader. No warranty is made as to the accuracy, completeness, or current applicability of the content; readers should not rely on it as a substitute for individualized professional advice and should consult qualified legal counsel licensed in their jurisdiction regarding any specific question. The views expressed are the author's own and do not represent the views of any employer, client, or affiliated organization. This document does not recommend for or against the substantive merits of any particular consensus change proposal, including BIP-110; it evaluates the readiness of activation processes. The model policy language in Appendix B is illustrative drafting for adaptation by qualified counsel and likewise does not constitute legal advice.

Acknowledgments

Thanks to the operators, miners, and users who keep this network running. To the dedicated volunteers who refuse to cut corners and build Bitcoin—line by line, meetup by meetup, and argument by argument. To Murch and Jameson Lopp, for taking time with an early draft. And to Ren Crypto Fish, Steve Lee, and Lyn Alden, whose *Analyzing Bitcoin Consensus* mapped the descriptive terrain this paper’s normative standards depend on.

Per the disclosure standard this paper proposes (§3.3): Murch is a BIP editor and the author of BIP-3, and Jameson Lopp is a co-author of BIP-361—both documents discussed in these pages—and §2.2 quotes Lopp’s 2026 block-size retrospective. Their time with early drafts implies no endorsement of this framework or its scoring, and the discussions of their work were drafted independently.

—Asaf

Contents

Acknowledgments	v
Abstract	ix
Summary for Policymakers	xi
1. The Problem	1
1.0 How to Read This Paper	1
1.1 The Absence of Standards	1
1.2 The Inscription Era and the Wave of Restrictive Proposals	2
1.3 BIP-110 as Case Study	4
1.4 The Stakes	6
1.5 Relation to Prior Work	7
2. Historical Precedent	13
2.1 P2SH (BIP-16) — 2012	13
2.2 The Block Size Wars (2015–2017)	14
2.3 SegWit (BIP-141) — 2017	15
2.4 Taproot (BIP-340/341/342) — 2021	15
2.5 Summary of Activation Parameters	16
3. The Framework	19
3.1 Proposal Submission Requirements	19
3.2 Minimum Review Period	20
3.3 Code Audit Requirements	21
3.4 Activation Threshold Standards	23
3.5 Chain Split Risk Assessment	29
3.6 Sunset and Reversibility Requirements	30
3.7 Hard Fork Subtypes: Scheduled vs. Emergency	31
4. Legal Analysis	35
4.0 Purpose of the Legal Analysis	35
4.1 Negligence	36
4.2 Tortious Interference	41
4.3 Fiduciary Duties	42

4.4	Mining and Node Operator Liability	43
4.5	Regulatory Consequences	44
4.6	Comparative Note: Common-Law Jurisdictions and EU Software Liability . .	45
5.	Proposed Standards	49
5.0	Red Flags: Is This Proposal Risky?	49
5.1	The Consensus Change Readiness Checklist	50
5.2	Scoring	54
5.3	On the Measurability of the Criteria	58
5.4	Worked Examples: Taproot, BIP-110, SegWit2x, and a Neutrality Test	62
6.	Objections and Responses	69
6.1	“Bitcoin has no governance.”	69
6.2	“Anyone can run whatever software they want.”	69
6.3	“This framework would prevent necessary changes.”	69
6.4	“Who decides whether the standards are met?”	70
6.5	“The legal analysis is speculative.”	71
6.6	“If the standards are not enforceable, what does the framework add?”	71
6.7	“This framework is itself a centralizing instrument.”	73
6.8	“This framework is a chilling instrument that will deter good-faith development.”	73
7.	Conclusion	77
	Appendix A — Documentary Record for the BIP-110 Case Study	79
	Appendix B — Adoption Kit: Model Policy Language	81
	Glossary of Technical Terms	83
	References	87

Abstract

At 55% hashrate enforcement, the chance of a six-block reorganization during activation is roughly **thirty percent**. By 90% or above it is **effectively zero** — five to seven orders of magnitude lower. The catastrophic failure mode of a consensus change is **computable**, and much of the governance argument is a fight over that one number.

Bitcoin has no formal process for evaluating proposed changes to its consensus rules. The Bitcoin Improvement Proposal (BIP) system provides a way to propose changes, but sets no minimum standards a proposal must meet before the community considers activation. There are no required review periods, no mandatory code audit standards, no agreed-upon activation thresholds, no chain split risk assessment methodology, and no framework for evaluating the legal and economic consequences of a failed activation.

This absence of standards has produced predictable results. The block size wars of 2015–2017 consumed years of developer time, fractured the community, and produced a contentious hard fork. The SegWit activation is widely credited to the threat of a novel User Activated Soft Fork (UASF) — BIP-148 — breaking the deadlock that had persisted after BIP-9 signaling failed to reach its 95% threshold over the preceding eight months. More recently, BIP-110 — a proposed temporary soft fork to restrict arbitrary data in Bitcoin transactions — was released with an activation client that multiple developers reported contained significant bugs. Its deployment paired that client with a 55% activation threshold far below any successful modern deployment and a six-week timeline from initial proposal to activation client — shorter than the review period of every modern successful soft fork, for SegWit and Taproot by more than an order of magnitude. The activation client was first distributed alongside stable Knots releases on node management platforms with no risk disclosure or visual differentiation, then bundled into the default release stream itself — so that the routine upgrade ladder terminates in a confirmation dialog that presents consensus alteration as the path forward.

This paper proposes a framework for evaluating Bitcoin consensus change proposals, and it begins from a claim these debates rarely confront: the catastrophic failure mode of a consensus change — a persistent chain split — has a computable risk profile, and that risk depends almost entirely on one variable, the share of hashrate that will enforce the new rules at activation. Section 3 models it and shows it is not a smooth function of that share but a cliff: at 55% enforcement the chance of a six-block reorganization during activation is roughly thirty percent; by 90% or above it is effectively zero, five to seven orders of magnitude lower. Much of the governance argument turns out to be a fight over a number — one the standard mathematics of chain-race security already supplies, but that the activation debate has almost never computed. The framework builds on that quantification, on the history of Bitcoin's

prior consensus changes, on the maturity-gate principles the open-standards tradition has applied to protocol governance for thirty years, and on a legal analysis of the liabilities an inadequately reviewed activation creates. It is practical, concrete, and immediately applicable. It proposes no changes to Bitcoin’s code; it proposes standards for the process by which such changes are evaluated, debated, and either adopted or rejected.

The author is a practicing litigator, solo Bitcoin miner, full node operator, and computer scientist — roles that carry their own operational stake in how block space is used. This framework is written from the intersection of those disciplines, and is deliberately limited to procedural standards so that its conclusions do not turn on the author’s view of the underlying inscription debate. The problems it addresses — governance, liability, technical risk, and economic consequence — cannot be fully analyzed from any single perspective.

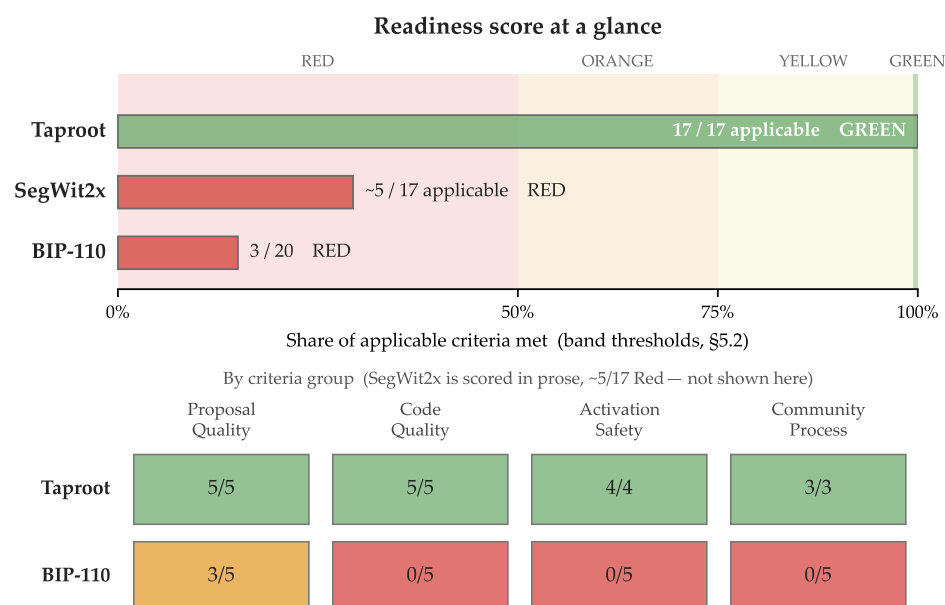
Summary for Policymakers

A summary for policymakers, operators, and time-pressed readers. It summarizes; the paper, its authorities, and its caveats govern.

For a policymaker, the headline is this: the risk of the worst outcome of mishandling a Bitcoin consensus change—a permanent split into two incompatible chains—is **computable** in advance, and it turns almost entirely on a single variable: the share of mining power that will enforce the change at activation. The gap is stark—roughly a thirty-percent chance of a damaging reorganization when only 55% enforce, falling to effectively zero by 90%. This framework is a way to measure that risk, and the broader readiness of any proposed change, before it ships rather than after.

What this is. A neutral, public **20-criterion readiness checklist** for judging whether a proposed change to Bitcoin’s shared consensus rules has been adequately reviewed before it goes live. The criteria fall in four groups—Proposal Quality, Code Quality, Activation Safety, and Community Process—and a proposal is scored Green / Yellow / Orange / Red on the share it meets. It is **not** a proposed law and **not** a change to Bitcoin’s code; it sets standards for the *process* by which changes are evaluated, debated, and adopted or rejected. It is procedurally neutral—the same standards apply whether a change would liberalize Bitcoin’s policy or restrict it.

What it found.



Applied to real proposals: **Taproot scores 17/17 (Green)**; **BIP-110 scores 3/20 (Red)**; and **SegWit2x scores roughly 5/17 (Red)**. To show the framework measures *process, not preference*, it also scores a hypothetical change the author's own mining revenue would favor — and that change scores Red too.

Why a policymaker should care. A rushed change that ignores documented red flags is not just a technical risk — it is a **legal** one. The framework's force is **deterrence, not enforcement**: a documented Red score is a dated, public record that supplies the *notice* and *foreseeability* the negligence and negligent-misrepresentation theories of the legal analysis (Section 4) depend on, should a change be activated anyway and cause loss. The flip side is the protection: **compliance is the defense; the score is the exposure.**

Does it require operators to coordinate? No. The framework **binds no one and asks for no agreement**: it publishes criteria each operator applies *unilaterally* to the listing, custody, and risk decisions it already makes. Independent parallel conduct on published criteria is not a Sherman Act violation; only an actual *agreement* to refuse would raise the antitrust concern — the posture the framework is designed never to form, and the model adoption language (Appendix B) is drafted accordingly.

Chapter 1

The Problem

1.0 How to Read This Paper

The framework proposed here is not an attempt to centralize authority over Bitcoin consensus. It is an attempt to give Bitcoin’s existing decentralized governance process a shared vocabulary for evaluating proposals.

The framework evaluates the ecosystem behavior surrounding a consensus-change proposal — the activation clients built around it, the review process applied to its code, the support evidence assembled in its favor, and the operational coordination before any deployment. It does not evaluate, and cannot evaluate, the editorial status of any BIP document. A BIP is a proposal; this framework asks which proposals merit activation, not which proposals merit filing.

Readers concerned with the question of authority itself — who decides? what makes one threshold “wrong” and another “right”? — should read §6.2 and §6.4 before §3. The framework is a tool, not a rule; if that distinction does not hold for a given reader, the standards that follow will read as something they are not.

Different sections serve different audiences. Protocol developers will find §3 and §5 most directly applicable. Businesses, exchanges, and custodians will find §3.5 and §4 most operationally relevant. Node operators and users will find §6 most useful for evaluating what the framework asks of them and what it does not.

If you read only two things, read §3.4 — where the chain-split risk turns out to be a computable cliff, not a slope (Figure 3.1) — and the worked examples at the head of §5.4 (Figure 5.2), which score Taproot, BIP-110, SegWit2x, and a change the author’s own mining would profit from against the same criteria. Technical terms are defined in the Glossary at the back (p. 83); a reader who hits an unfamiliar one can look it up there.

1.1 The Absence of Standards

Bitcoin’s consensus rules govern the creation, transfer, and validation of an asset with a market capitalization exceeding one trillion dollars — among the most consequential bodies of code in finance. Changes to these rules affect every participant in the network: miners who invest capital in hardware, node operators who validate transactions, developers who build applications, businesses that accept payment, and individuals who store wealth.

Despite these stakes, there is no formal standard for how changes to consensus rules should be proposed, evaluated, reviewed, tested, activated, or — critically — rolled back if they fail. The BIP process — established in BIP-1, refined in BIP-2, and governed since late 2015 by BIP-3, which replaced BIP-2 — provides a template for writing proposals and a taxonomy of proposal types. BIP-3’s revision deliberately narrowed the judgment calls given to BIP editors, making explicit what the repository’s design always implied: it is a filing system for proposals, not a gatekeeper of activations. Neither the process document nor the repository establishes minimum standards for activation safety, mandatory review periods, code quality requirements, or chain split risk assessment.

The result is an ad hoc system in which each consensus change proposal invents its own activation mechanism, sets its own threshold, defines its own timeline, and is evaluated by the community with no consistent framework. Some proposals receive years of careful review. Others are pushed to activation within weeks. The difference between these outcomes is decided not by any institutional process but by the personalities, politics, and persuasive abilities of the participants.

1.2 The Inscription Era and the Wave of Restrictive Proposals

In late 2022, the developer Casey Rodarmor made the first inscription on Bitcoin mainnet (December 14, 2022) and, in January 2023, publicly released the Ordinals protocol — a method for encoding arbitrary data into Bitcoin transactions by using Taproot’s tapscript capacity. The protocol enabled what came to be called “inscriptions”: image, text, document, and binary files embedded permanently within the witness data of Bitcoin transactions. After the protocol’s early-2023 release, inscription activity grew through 2023 to take up a substantial share of block space across multiple weeks of high demand.

The mechanism is technically simple but consequential for governance. Inscriptions place arbitrary bytes within Taproot’s script-path spend by enclosing them in an unexecuted `OP_FALSE OP_IF ... OP_ENDIF` envelope. Bitcoin nodes do not validate or execute the data within the envelope; the data becomes a permanent part of the chain’s witness record. The SegWit witness-data fee discount, which applies to the tapscript carrying the envelope, makes this storage method substantially cheaper, per byte, than alternatives such as `OP_RETURN`. Taproot (BIPs 340–342), activated in November 2021, introduced a new output type supporting Schnorr signatures, MAST-style script trees, and a clean soft-fork upgrade hook (the `OP_SUCCESS` opcodes), and its Tapscript rules lifted the legacy per-script size and opcode limits — all within SegWit’s existing witness discount and weight budget, raising no ceiling. That new spend path, combined with SegWit’s pre-existing witness discount, also supported a use case its designers did not anticipate.

The inscription debate became, by mid-2023, the most contentious sustained policy disagreement in Bitcoin since the block size wars, and it has continued through 2026. Two camps emerged. Proponents argued that inscriptions pay full fees, displace no monetary transac-

tions in equilibrium, and represent a legitimate use of the open block space markets have priced. Opponents argued that inscriptions consume scarce block space and witness storage with non-monetary content, raise long-term storage and bandwidth costs for every full-node operator, and degrade Bitcoin's function as monetary infrastructure. Both arguments draw on real considerations; neither settles the question. Reasonable participants have reached opposite conclusions and continue to debate the question in good faith.

This paper does not evaluate the substantive merits of that debate. The question of what counts as a legitimate use of Bitcoin's base layer is for the network's stakeholders to resolve through the iterated coordination processes *BCAP* describes — not for any single framework, including this one, to settle. What this paper does evaluate is the procedural conduct surrounding proposals that would resolve the question by consensus change. Beginning in 2023, multiple proposals emerged seeking to restrict the methods that enable inscriptions: tightening data-embedding limits, restricting tapscript witness contents, or constraining specific transaction patterns. BIP-110 — the Reduced Data Temporary Softfork analyzed in §1.3 — is the most-developed instance of this wave and the one whose activation client reached release.

The proximate trigger for the wave's culminating escalation was a 2025 change at the level of relay policy, not consensus. Bitcoin Core's v30 release (October 2025) raised its long-standing default relay limit on OP_RETURN data carriage and permitted multiple data outputs per transaction. The change touched only standardness and relay policy — not the consensus rules — but data-restriction advocates read it as Core endorsing arbitrary-data use, and it precipitated a migration of node operators toward Bitcoin Knots, which retained the stricter default. BIP-110 emerged from that schism. The dispute the framework evaluates is therefore not abstract: it is the live Core-versus-Knots disagreement over Bitcoin's data-carrying policy, which BIP-110 escalated into a proposed consensus change.

The framework presented in this paper is procedurally neutral on the underlying disagreement. The same evaluation applies whether a proposal would liberalize Bitcoin's policy space or restrict it. The same standards apply whether the proponents are inscription advocates or inscription critics. A consensus change pushed through inadequate review, low activation thresholds, and conflict-of-interest dynamics raises the same governance and legal exposure no matter which side of the substantive debate it sits on. The framework asks one question of each proposal: is it ready, by the standards of Section 3 and Section 5, for the community to engage with its activation? That question is separate from whether the underlying change is desirable on the merits.

1.3 BIP-110 as Case Study

BIP-110—the Reduced Data Temporary Softfork—implicates most of the failure modes this framework is designed to surface, and is examined here as a case study. Originally proposed as BIP-444 in late October 2025, the proposal sought to restrict methods of embedding arbitrary data in Bitcoin transactions. Its stated goal was to protect Bitcoin’s function as monetary infrastructure by limiting what proponents called “spam” uses of block space.

The proposal’s technical merits are debatable. Reasonable people disagree about whether inscriptions, ordinals, and large OP_RETURN payloads are a legitimate use of Bitcoin’s base layer or an abuse of shared block space—a characterization each side states in its own terms. This paper takes no position on that question. The problems with BIP-110 are procedural, not substantive:

A. Activation threshold. BIP-110 specified a 55% miner signaling threshold for a User Activated Soft Fork. More precisely, BIP-110 operates as a LOT=true mechanism: an early lock-in trigger at 55% signaling and a mandatory-signaling period around August 2026 (blocks 961,632–963,647) in which BIP-110-enforcing nodes reject all non-signaling blocks, guaranteeing lock-in, with activation at `max_activation_height` block 965,664 (approximately September 2026). The framework’s position on LOT=true is set out in §3.4. This 55% threshold is dramatically lower than every successful modern precedent. SegWit’s BIP-9 deployment required 95% miner signaling; when that stalled, BIP-91 created a parallel mechanism at 80%, and the BIP-148 UASF threatened to reject non-signaling blocks entirely. Taproot activated at 90% via Speedy Trial. A 55% signaling threshold provides no assurance that the share of hashrate enforcing the new rules at activation will exceed 55%; the rest may keep producing blocks valid under the legacy rules but invalid under the new ones. Signaling at lock-in is not equivalent to enforcement at activation, and the divergence between the two is precisely the mechanism by which low-threshold soft forks produce persistent minority chains. This is not a theoretical risk: at 55% enforcement the framework’s model (§3.4) puts the probability of a six-block reorganization during activation near thirty percent.

B. Code quality. The activation client, released in late 2025 as a fork of Bitcoin Knots, was found to contain significant bugs. Multiple Bitcoin developers reported that the client’s test suite—including activation tests its own author had written—was failing, such that its activation behavior could not be verified, and warned that users running the code might accidentally fork themselves off the network. Public commentary from reviewers raised concerns about the code’s structure and quality; whatever tools or methods were used to produce it, the activation client did not receive the level of independent review that consensus-critical software demands.

C. Review period. From initial proposal to release of the activation client, BIP-110 moved through the pipeline in approximately six weeks—from the initial bitcoin-dev mailing list post on 26 October 2025 to release of the first activation client (v0.1rc1) on 10 December 2025.

By comparison, SegWit was proposed in December 2015 and did not activate until August 2017—a twenty-month process. Taproot was first proposed in January 2018 and activated in November 2021—nearly four years. Six weeks is not a review period. It is a rush to deployment.

D. Activation client distribution. On at least one node management platform, the BIP-110 activation client was initially listed as a selectable version option in the same dropdown menu as stable Knots releases, with no warning label, risk disclosure, or visual differentiation. Selection was deliberate—a node operator had to affirmatively choose the BIP-110 version—but the presentation treated consensus-altering software the same as routine maintenance releases. A node operator who understood version management but not the implications of BIP-110 specifically could have activated consensus-changing code believing it was a standard update. The absence of any risk disclosure at the point of selection was the governance failure, not the availability of the option itself.

A subsequent escalation merged the activation code into the project’s default release stream. As of `v29.3.knots20260508` (released 9 May 2026), the stable Bitcoin Knots release includes the BIP-110 (RDTs) activation rules and prompts the operator to confirm activation at runtime, via either a configuration directive (`consensusrules=rdts`) or a GUI dialog. The non-RDTs variant remains available—as the one-day-older `v29.3.knots20260507` build—labeled as the discouraged option. The confirmation is real; the operator must affirmatively accept the consensus change. The governance failure is now structural rather than presentational: the routine upgrade ladder, which an operator follows simply to receive ordinary bug fixes and feature improvements, ends in a confirmation dialog that presents consensus alteration as the path forward. The release notes characterize the change as fixing “critical vulnerabilities in long-standing network design” and assert that the upgrade “already has broad community support”—an assertion made while voluntary block-signaling for the change stood below one percent, and one the legal analysis returns to (§4.1). Whether or not one accepts the vulnerability characterization—a question that turns on the substantive inscription debate this paper does not enter—the governance failure is routing a consensus change through the routine-upgrade confirmation flow: declining the change means giving up the maintenance benefits of the newer release, while accepting means defaulting into a consensus alteration by the path of least resistance.

E. Sunset mechanism. BIP-110 includes an automatic expiry at a defined block height, after which the new rules cease to be enforced. This is a meaningful improvement over proposals that lack any deactivation mechanism. However, there is no public evidence that the sunset mechanism was tested on testnet to confirm that the transition back to pre-activation rules would occur without consensus failures—a concern amplified by the significant bugs found in the activation client itself. A sunset clause that has not been demonstrably tested is a promise, not a guarantee. Further, the one-year duration was chosen without empirical justification for why one year is the right period, and the proposal contained no defined process for evaluation at the end of the enforcement period.

The “BIP-110” addressed throughout this paper refers to the activation attempt — the RDTs client, the surrounding ecosystem campaign, and the operational coordination assembled around the proposal — not to the BIP-110 document filed in the BIPs repository. A BIP is a proposal; the framework’s evaluative scope is the ecosystem behavior that would convert a proposal into network state. BIP-361 (*Post Quantum Migration and Legacy Signature Sunset*, Lopp et al. 2026), assigned its number in February 2026, is an instructive contemporaneous example: a proposal of consequential scope and ongoing community debate, with (as of mid-2026) no implementation code and no activation campaign — a BIP, in the editorial sense, without any of the ecosystem facts this framework evaluates.

The question of how any single threshold could be “set” in a network with no central authority is foundational and is addressed in §6.2 and §6.4. The short answer: it cannot, in any binding sense. BIP-110’s proponents proposed 55%; whether the network adopts it is being decided by tens of thousands of independent operator decisions, which through July 2026 had produced voluntary block-signaling of well under one percent — far short of the 55% lock-in threshold (the figure is the peak observed share of blocks signaling readiness during the voluntary period). Because BIP-110 deploys as a LOT=true mechanism, the voluntary period’s verdict does not end the deployment: mandatory signaling is scheduled to open at block 961,632 (early August 2026) and lock in at block 963,648 (late August 2026), with forced activation at block 965,664 (early September 2026); whether that forced activation yields enforcement or a persistent minority chain will turn on the economic-node response the framework addresses in §5.2.

1.4 The Stakes

A failed consensus change activation is not a software bug that can be patched. It is a potential fracture of the monetary network. When a chain split occurs without replay protection, transactions valid on one chain may be valid on the other. Users can lose funds. Exchanges must choose which chain to list. Contracts denominated in Bitcoin become ambiguous. The economic damage is real, quantifiable, and potentially irreversible.

The damage is not uniform across everything the ledger carries, and the gradient matters. Bitcoin’s chain today carries more than balances — it carries inscribed art, backup records, and timestamp anchors whose whole job is to prove a document existed by a date. The non-monetary uses survive a split better than money does: the full pre-fork history sits on both chains, and both keep accepting data afterward. But surviving is not coming through whole. Fungible value breaks outright — every balance doubles into two incompatible histories, and the replay and listing problems above land on every holder at once. Data does not break; it gets harder to use, because every reference to “the Bitcoin blockchain” now names two candidates instead of one. Nothing the chain carries comes out of a split better off. That is why this paper’s worst-case analysis runs on the economic-ledger function alone: money is the use a split damages most, and a standard strict enough to protect it protects every harder use with it.

One concrete case shows what “harder to use” means in practice, and it reaches courtrooms. A Bitcoin timestamp proof — an OpenTimestamps-style anchor that commits a document’s hash to the chain — proves the document existed by a given date without trusting any server or custodian. The proof works because “the Bitcoin blockchain” names one canonical chain. After a split it names two, and every proof invites a threshold dispute over which chain’s history counts — litigated before the tribunal ever reaches what the proof proves, the evidence-law face of the contract ambiguity §4.2 maps. A split also weakens the proofs themselves, and it does so asymmetrically. An anchor placed before the split lives in the history both chains share, so erasing it means rewriting both. An anchor placed after the split lands on a single chain — a weakened one, because hashrate is one number doing two jobs: before activation it decides whether the network splits, and after a split it decides how expensive each half’s history is to rewrite. The minority side of Ethereum’s 2016 split ran that experiment: Ethereum Classic was 51%-attacked three times in August 2020 alone, the deepest attack reorganizing more than 7,000 blocks. The author’s interest belongs on the table here, as it does for mining in §5.4: he is building Bitcoin-anchored evidence timestamping into legal case software, so an unambiguous, expensive-to-rewrite chain is worth money to him. The mining interest runs against this paper’s conclusions; this one runs with them. Both are disclosed on the same principle: a framework that scores conflicts of interest should not conceal its author’s.

The most damaging chain split without replay protection occurred in March 2013, when a database incompatibility between Bitcoin versions 0.7 and 0.8 caused a six-hour fork that included a successful double-spend attack (the technical post-mortem is BIP-50); a shallower accidental fork — six blocks — followed BIP-66’s activation in July 2015, when miners built on an invalid block without validating it. The 2017 SegWit2x proposal came within days of producing another before being called off. The Bitcoin community has been fortunate. Fortune is not a governance strategy.

1.5 Relation to Prior Work

“Rough consensus and running code.”

— David Clark (1992), as recorded in RFC 7282; the IETF credo this framework’s maturity gates descend from

The framework proposed in this paper builds on, and is intended to be read alongside, the most developed prior analysis of Bitcoin’s consensus-change dynamics: Ren Crypto Fish, Steve Lee & Lyn Alden, *Analyzing Bitcoin Consensus: Risks in Protocol Upgrades* (Nov. 2024) [hereinafter *BCAP*]. *BCAP* categorizes participants in consensus changes into six stakeholder groups — Economic Nodes, Investors, Media Influencers, Miners, Protocol Developers, and Users/Application Developers — and analyzes how the relative power of each group fluctu-

ates across the phases of a proposal’s lifecycle. It introduces a “State of Mind” framework distinguishing degrees of stakeholder engagement, develops a detailed scenario analysis of consensus changes deployed through alternative clients, and concludes with thirteen evaluation questions and twelve indicators stakeholders may use to assess proposals in real time. *BCAP*’s draft was reviewed by a substantial cross-section of Bitcoin’s protocol-development and analytical community.

BCAP and the present framework address the same problem from different vantages. *BCAP* is descriptive: it explains how Bitcoin consensus emerges from the iterated interactions of stakeholders with shifting powers and divergent incentives. This paper is normative: it proposes the minimum standards a proposal should meet before stakeholders rationally engage with its activation. The two works converge on several substantive points. Both treat sub-overwhelming Economic Node adoption as the central failure mode in contested soft-fork scenarios — *BCAP* through its bounty-claim and chain-split risk analysis (*BCAP* §3.5.2), this paper through its activation-threshold standards (§3.4) and chain-split risk-assessment requirement (§3.5). Both treat miner signaling as a necessary but insufficient indicator of community consensus. Both reject naïve majoritarianism as a consensus-determination heuristic. And both identify the independent economic-node response as the operative remedy when an activation proceeds without genuine consensus: *BCAP* describes it descriptively, as the mechanism by which markets re-equilibrate after a contested split; this paper prescribes it normatively, as the recommended response to a proposal classified as Red under §5.2.

This paper extends *BCAP*’s analytical framework in four respects. First, it turns qualitative recommendations into numerical floors: minimum activation thresholds (90% MASF, with sub-80% presumptively dangerous and sub-60% reckless); minimum review periods tied to risk category (twelve months for moderate-risk soft forks, twenty-four months for high-risk soft forks, thirty-six for hard forks, refined by subtype in §3.7); and minimum testnet deployment duration (three months). Second, it quantifies the relationship between activation threshold and chain-split exposure: §3.4 models post-activation hashrate as a Bernoulli process and derives concrete reorganization probabilities — approximately thirty percent at $E = 0.55$ for a six-block deficit, falling to roughly 2×10^{-8} at $E = 0.95$. This quantification is consistent with *BCAP*’s qualitative conclusion that low and high Economic Node adoption produce categorically different risk profiles, but supplies the order-of-magnitude estimates that the qualitative analysis leaves open. Third, it develops a legal-liability framework — negligence (with explicit attention to the economic-loss rule and the available routes around it), tortious interference, fiduciary duty, mining-pool contract obligations, and regulatory consequences. *BCAP* does not address legal exposure; this paper offers it as a distinct analytical layer whose conclusions are relevant to every stakeholder group *BCAP* identifies. Fourth, it consolidates the above into a binary twenty-criterion scoring rubric with classification bands (§5), enabling structured evaluation in place of free-form weighing of considerations.

Several elements of *BCAP* the present paper does not duplicate. The stakeholder taxonomy and the power-over-time analysis (*BCAP* §§3.2, 3.3.2), the State of Mind framework (*BCAP* §3.1), the investor-segment analysis distinguishing self-custodying holders from institutional,

corporate-treasury, and exchange-traded-fund segments (*BCAP* §3.2.2), and the alternative-client adoption-pathway analysis (*BCAP* §3.5.1) all remain the more developed treatments of their subjects. This paper refers readers to *BCAP* for those questions. The two frameworks are intended to function together: *BCAP* supplies the theory of how stakeholders shape consensus; this paper supplies the standards by which stakeholders may judge whether a particular proposal is ready for that process to begin.

The minimum-review-period floors proposed in §3.2 are grounded in empirical observation rather than aesthetic preference. Jameson Lopp’s longitudinal analysis of Bitcoin Core node-software adoption documents that reaching 95% node-operator adoption of a Core release took roughly a year before 2018 and has since approximately doubled to two years — possibly three for the most recent releases — with peak adoption now lagging six to nine months behind release. See Jameson Lopp, *When Do Bitcoin Node Operators Upgrade?*, blog.lope.net/when-do-bitcoin-node-operators-upgrade/. A twelve-month minimum review period for a moderate-risk soft fork spans roughly half a current upgrade cycle, and a twenty-four-month minimum for a high-risk soft fork roughly one full cycle; both are set deliberately as floors, not as the adoption target. These floors are anchored, not arbitrary, but — like the ninety-percent threshold (§3.4) — they are floors, not ceilings: the same upgrade data would comfortably support longer review periods, and a stakeholder is free to demand more. They are the minimum the network’s own observed upgrade dynamics justify for activation signaling to begin from a position of broad enforcement readiness rather than speculative anticipation.

The legal analysis in Section 4 engages a smaller but distinct body of prior work. Angela Walch’s argument that core protocol developers exercise discretionary authority over property interests sufficient to trigger fiduciary obligations is the central academic contribution to the question and is engaged directly in the fiduciary-duty analysis below. The Tulip Trading litigation (*Tulip Trading Ltd v van der Laan* [2023] EWCA Civ 83) is the most developed common-law treatment and frames the standard against which a developer’s conduct would be evaluated. The other legal questions raised by chain splits — contract interpretation of “Bitcoin,” the tax treatment of forked coins under IRS Revenue Ruling 2019-24, exchange custodial obligations during a fork — have received episodic treatment but no consolidated analysis. The legal-analysis section of this paper consolidates that scattered authority.

This paper occupies the operational and legal layer of an emerging body of work on blockchain governance — a literature that includes De Filippi and Wright on code as a mode of regulation, the on-chain-versus-off-chain governance debate (Buterin; Zamfir), and the textbook account of off-chain coordination in Narayanan and colleagues. This framework takes a side within that debate. With Zamfir and Buterin — both skeptics of binding on-chain governance, though Buterin allows coin-voting a limited, non-exclusive role — it locates legitimate consensus in off-chain coordination rather than a tallied vote; against the purely descriptive accounts, it adds what they leave out, a published standard against which that off-chain process can be measured and, when it is bypassed, held to account. Where De Filippi and Wright read code as a mode of regulation, this paper treats the *activation* of code as the regulated act, and

locates accountability not in the code but in the documented process that ships it. It is most useful when read alongside *BCAP*, not in place of it.

Governance models in adjacent open protocols

The framework’s central move — no activation until running, independently reviewed code has cleared minimum maturity gates — is not new to open-protocol governance, though it is new to Bitcoin consensus change. It is the operating premise of the Internet’s own standards process. RFC 7282 states the IETF credo as “rough consensus and running code,” an expressly anti-majoritarian formulation that, like this framework and like *BCAP*, rejects head-counting as a substitute for technical readiness. RFC 2026, the Internet Standards Process, makes the maturity bar concrete and tiered: a specification advances to Draft Standard only on “at least two independent and interoperable implementations from different code bases” with “sufficient successful operational experience,” and the top Internet Standard level is reserved for one with “multiple, independent, and interoperable implementations with substantial operational experience.” (RFC 6410 (2011) later consolidated the track to two maturity levels; the two-implementation bar survived at the Internet Standard level.) That is the running-reviewed-code-before-activation gate restated in the vocabulary of the body that invented it.

The lineage is documented, not inferred. EIP-1 records that it is “derived heavily from Bitcoin’s BIP-0001 written by Amir Taaki which in turn was derived from Python’s PEP-0001” — a PEP-1 to BIP-1 to EIP-1 descent tracing back to the Internet’s request-for-comments practice, which the IETF inherited. Bitcoin’s proposal system is a branch of that tree, and the standards from which the present framework’s gates descend are its forerunners. The contrast at the other end of that lineage is instructive: Ethereum runs a frequent-coordinated-upgrade cadence through its All Core Devs calls, a process built to ship hard forks on a regular schedule, whereas Bitcoin’s conservatism leaves it without any comparable maturity discipline — which is the gap this framework fills.

What the open-standards tradition supplies is the maturity-gate insight: no central authority, but a refusal to treat unimplemented or single-implementation code as ready. What it does not supply is the rest of this paper. The IETF process quantifies no activation-threshold risk and assigns no legal consequence to shipping past a documented defect. This framework brings the maturity gate into the Bitcoin consensus-change setting and adds two layers those models lack — a computed chain-split probability tied to the enforcement share (§3.4) and a tort and contract liability analysis (§4) — layers that make the gate enforceable in foreseeability terms rather than merely customary. Measured against the open-standards maturity bars, BIP-110 fails at the threshold: an activation client whose own tests were reported failing and no second, independent implementation. RFC 2026 would not advance a specification on that record; neither should a consensus change.

Chapter 1 in brief

Bitcoin has a process for filing proposals (the BIP system) and no process for evaluating them. Every consensus change reinvents its own playbook for review, activation, and rollback. The block size wars, SegWit's near-miss in 2017, and the inscription-era wave of restrictive proposals culminating in BIP-110's six-week sprint to deployment are all symptoms of the same gap.

This paper proposes minimum standards a proposal should meet before stakeholders rationally engage with its activation. It is meant to be read alongside *BCAP* (Crypto Fish, Lee, Alden, 2024), which describes *how* consensus emerges from stakeholder interaction; this paper proposes *what* a proposal should look like to deserve that engagement. Different vantages on the same problem.

Chapter 2

Historical Precedent

Bitcoin has made many consensus changes since 2009. The most significant ones offer useful precedent for setting governance standards.

2.1 P2SH (BIP-16) — 2012

Pay-to-Script-Hash was one of Bitcoin's first contentious soft forks, and it is the most instructive precedent this framework has — the one time Bitcoin ran a roughly 55% threshold for a real consensus change, the same threshold BIP-110 later adopted. It began as a design dispute: after the original OP_EVAL approach (BIP-12) was abandoned over a security flaw, Gavin Andresen's BIP-16 carried the developer consensus while Luke Dashjr's competing BIP-17 (OP_CHECKHASHVERIFY) remained the minority alternative. Activation was by coinbase signaling — if 550 of the preceding roughly 1,000 blocks (about 55%) carried the string /P2SH/, the new rules would take effect.

The threshold did not carry it. By the February 2012 polling date fewer than half of the last 1,000 blocks were signaling, so the planned March 1 activation could not safely proceed. The partial enforcement then produced exactly the divergence the framework's model predicts: miners still enforcing the March 1 date rejected blocks from the majority that were not, and the network experienced a real, multi-block chain fork. The change activated only after the parameters were re-cut — a second poll in mid-March, then a hardcoded flag day at Unix timestamp 1,333,238,400 (April 1, 2012; block 173,805). The 55% signal did not activate P2SH; a flag day did, after the signal fell short and the chain forked.

Lesson: The one time Bitcoin ran a 55% threshold for a real consensus change, the signal failed on the first attempt and the chain split — exactly as the random-walk model of §3.4 predicts. P2SH succeeded only because it was a small change to a network worth a tiny fraction of today's, rescued by a flag day rather than carried by the threshold. BIP-110 proposes the same 55% threshold on a far more consequential change to a trillion-dollar network. That is not precedent for safety; it is the precedent for the failure mode.

2.2 The Block Size Wars (2015–2017)

The block size debate consumed more community energy, developer time, and political capital than any other event in Bitcoin’s history. Multiple proposals competed: BIP-101 (8 MB blocks), BIP-102 (2 MB blocks), Bitcoin XT, Bitcoin Classic, Bitcoin Unlimited, and ultimately SegWit2x. The conflict produced the Bitcoin Cash hard fork in August 2017 and nearly produced a second split with SegWit2x in November 2017. The standard book-length history of the conflict is Jonathan Bier’s *The Blocksize War* (2021); this chapter draws the governance lessons and leaves the narrative history to it. Reviewing that book in mid-2026, Jameson Lopp — a participant who argued each side of the war at different points — distilled its answer to the question of who gets to change Bitcoin’s rules: “not miners, not companies, not famous developers, not meetings of industry leaders, but economically active users who choose what rules their nodes enforce and in the event of a split, which coin they choose to hold.” And he added a caution from personal experience about how much of such a war the public record captures: “there are many dynamics the book doesn’t cover, especially with regard to communications and strategizing in private channels. The public debates were only part of the battle for control over the future of Bitcoin.” See Jameson Lopp, *Reflections Upon Bitcoin’s Block Size Debate* (July 2026), blog.lope.net/reflections-upon-bitcoins-block-size-debate/. Both observations carry into this framework: the first is the economic-node authority its response mechanism rests on (§5.2); the second is a reason its criteria score only what a campaign publishes — the private half of an activation battle is real, and it is exactly the half no standard can verify.

The block size wars showed several principles any governance framework must account for:

Miner signaling is unreliable as a measure of community consensus. Mining pools signaled support for proposals their users did not endorse. The SegWit2x “New York Agreement” secured signatures from companies representing over 80% of hashrate, yet the proposal collapsed when it became clear that node operators and users would not follow.

Economic nodes matter more than hashrate. Exchanges, payment processors, and major holders decide which chain carries economic value. A chain with 90% of the hashrate but no exchange listings and no merchant adoption is worthless. Governance frameworks must account for economic consensus, not just miner signaling.

Hard forks are permanent and expensive. Bitcoin Cash continues to exist as a separate chain with a fraction of Bitcoin’s value. Every contested hard fork fragments the ecosystem, confuses users, and creates legal ambiguity about which chain constitutes “Bitcoin” for contractual and regulatory purposes.

2.3 SegWit (BIP-141) — 2017

Segregated Witness was proposed in December 2015 and activated in August 2017 via BIP-9 version bits signaling with a 95% activation threshold. When miner signaling stayed well below the 95% threshold through early 2017, the community developed BIP-148 — a User Activated Soft Fork that would have begun rejecting non-SegWit blocks on August 1, 2017, regardless of miner signaling.

The threat of BIP-148 — and the risk that it would cause a chain split — motivated miners to signal for SegWit. BIP-91 locked in around July 20–21, 2017, and from July 23 its enforcing miners rejected blocks that did not signal for BIP-141; every block was therefore signaling for SegWit before the August 1 UASF deadline, which satisfied BIP-148 nodes. BIP-141’s own 95% lock-in followed on August 8, 2017, with activation at block 481,824 on August 24. This episode established the UASF as a credible activation mechanism but also demonstrated its risks: had miners not capitulated, BIP-148 nodes would have split from the main chain.

Lesson: UASFs are a tool of last resort, not a standard activation mechanism. BIP-148 worked because SegWit had overwhelming community support and years of review. BIP-110 applied the same mechanism with weeks of review and signaling that never approached its threshold — the conditions §3.4 classifies as reckless.

2.4 Taproot (BIP-340/341/342) — 2021

Taproot is the strongest precedent for careful Bitcoin consensus-change governance. Its design was first floated on the bitcoin-dev mailing list in January 2018 and formally specified in BIPs 340, 341, and 342 in 2020; it then went through years of review, extensive formal analysis of its cryptographic primitives (Schnorr signatures), multiple rounds of community feedback, and a novel activation mechanism (Speedy Trial) that gave a defined three-month signaling window with a built-in timeout. Choosing Speedy Trial over BIP-8 with LOT=true and BIP-8 with LOT=false reflected a long bitcoin-dev mailing list debate over the right balance between activation speed, miner authority, and economic-node authority.

Taproot activated in November 2021 with 90% miner signaling within the Speedy Trial window. There was no chain split, no community fracture, and no economic disruption.

Lesson: A long review period, a high activation threshold, and a defined timeline with a built-in failure mode each contributed to Taproot’s clean activation. BIP-110 offered weeks of review, a 55% threshold, and a sunset clause that was never demonstrably tested.

2.5 Summary of Activation Parameters

Proposal	Threshold	Review Period	Mechanism	Outcome
P2SH (2012)	55%	~3 months	Coinbase signal + flag day	Activated (flag day)
SegWit (2017)	95%	20 months	BIP-9 + UASF	Activated
SegWit2x (2017)	80% (SegWit leg)	~6 months	Hard fork	Canceled
Taproot (2021)	90%	~4 years	Speedy Trial	Activated
BIP-110 (2025–2026)	55%	~6 weeks	UASF (Knots, LOT=true)	Forced activation ≈ Sep 2026 (voluntary < 1%)

Note: the percentage in BIP-110's Outcome entry is the peak observed share of blocks signaling readiness during the voluntary period. That 55% lock-in threshold was never approached during voluntary signaling (the peak share stayed under one percent through July 2026); because BIP-110 deploys as a LOT=true mechanism, mandatory signaling is scheduled for blocks 961,632–963,647 (August 2026), with forced activation at block 965,664 (early September 2026). The ~6-week figure is proposal-to-first-activation-client, not the signaling-window length. SegWit2x's 80% was btc1's signaling threshold for the SegWit leg; the 2x hard fork itself was unconditional. P2SH's 55% coinbase poll fell short of its threshold on the first attempt (February 2012) and the chain forked between enforcing and non-enforcing miners; it activated only on the April 1, 2012 flag day (block 173,805), not by the signal (§2.1).

The pattern is clear: successful consensus changes correlate with high activation thresholds, long review periods, and broad community buy-in. Failed or stalled proposals correlate with low thresholds, rushed timelines, and narrow support. This is not coincidence. It is the predictable result of governance dynamics any framework must formalize.

Figure 2.1 places these episodes on a common timeline.

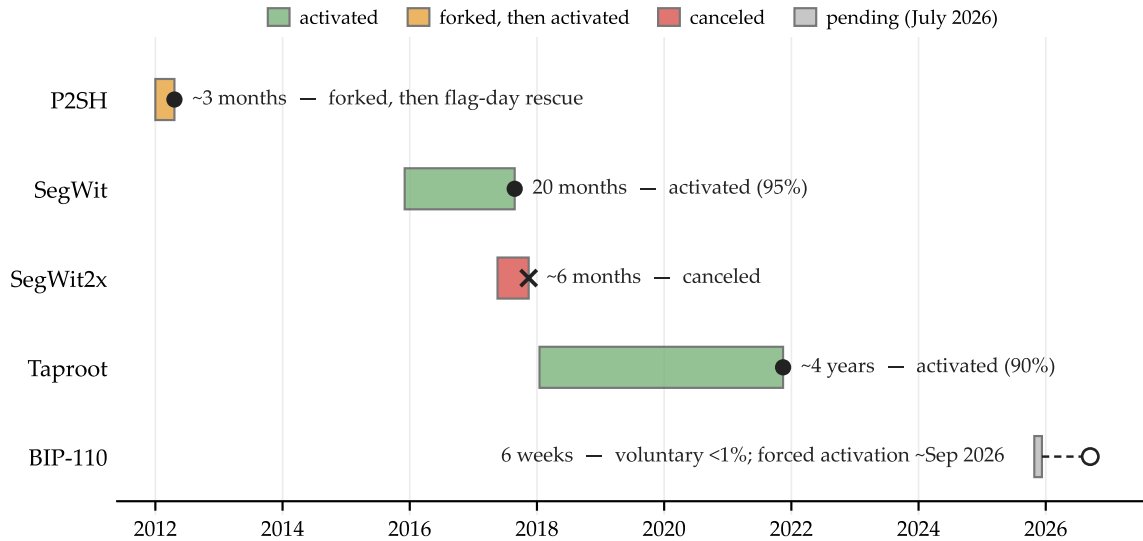


Figure 2.1: Review timelines of five consensus changes, 2012–2026. Each bar runs from first proposal to activation or cancellation—for BIP-110, to its activation-client release, with the dashed line marking the scheduled forced activation ($\approx$ Sep 2026)—drawn to scale; the comparison that matters is the bar’s length: twenty months for SegWit, about four years for Taproot, six weeks for BIP-110.

Chapter 2 in brief

Bitcoin’s prior consensus changes show one pattern. P2SH (2012) ran a 55% threshold, the signal fell short, and the chain forked; a flag day—not the threshold—carried it, and only because the change and the network were small. SegWit (2017) took twenty months and an unprecedented UASF threat to escape miner deadlock. Taproot (2021) used 90% signaling, a defined window, and years of review. The successful pattern: high threshold, long review, broad consensus. The failed and stalled patterns inverted at least one of those three.

Chapter 3

The Framework

The following framework establishes minimum standards for Bitcoin consensus change proposals. These standards are not meant to be enforced by code — Bitcoin has no central authority that could enforce them. They are meant to serve as a public benchmark the community can use to evaluate proposals. A proposal that meets these standards deserves serious consideration. A proposal that fails to meet them should be treated with appropriate skepticism.

3.1 Proposal Submission Requirements

A consensus change proposal submitted for community review should include, at minimum:

A. Problem Statement. A clear, specific description of the problem the proposal addresses, supported by empirical data where available. “Bitcoin should do one thing and do it well” is a philosophy, not a problem statement. “The UTXO set has grown by X% in Y months due to Z transaction type, imposing quantifiable costs of \$W per node operator per year” is a problem statement.

B. Technical Specification. A complete technical specification of the proposed change, including all modified consensus rules, their interaction with existing rules, and a formal analysis of edge cases. The specification should be detailed enough to permit independent implementation.

C. Backward Compatibility Analysis. A full analysis of the proposal’s impact on existing transactions, wallets, applications, and use cases. It should identify every category of transaction or script that would become invalid under the new rules and quantify, as far as possible, the number of affected users and the value at risk.

D. Activation Mechanism. A fully specified activation mechanism including signaling method, threshold, signaling window duration, timeout behavior, and defined failure mode. The activation mechanism should be described precisely enough to permit independent implementation and verification.

E. Rollback Procedure. For proposals self-described as “temporary,” a self-executing sunset mechanism is required (see §3.6). For other soft-fork proposals, no rollback procedure is

required, but the proposal should not affirmatively obstruct future reversal — the consensus rules should remain technically reversible via standard soft-fork or hard-fork mechanisms.

F. Reference Implementation. A complete, functional reference implementation against a current release of Bitcoin Core or a compatible client. The reference implementation must include a comprehensive test suite.

3.2 Minimum Review Period

No consensus change proposal should proceed to activation signaling until it has completed a minimum review period. How long that period should be depends on the scope and risk of the change. The categories below track a proposal’s risk to existing holdings and to network unity — whether it can invalidate currently valid transactions or split the chain — not the direction of the change: a proposal that loosened an existing limit in a way that endangered already-confirmed outputs would fall in the same high-risk category as one that tightened it. The framework’s risk taxonomy is neutral as between liberalizing and restricting changes.

Category 1: Low-Risk Changes

Changes that tighten existing policy without altering the consensus boundary — for example, reducing default mempool relay limits. These are not consensus changes and do not require this framework. Individual node operators can adopt or reject them at will.

Category 2: Moderate-Risk Consensus Changes

Soft forks that add new validation rules without invalidating any currently valid transaction type. Examples include Taproot and SegWit, which added new script capabilities by tightening rules over spend paths no deployed wallet used. **Minimum review period: twelve months** from publication of a complete technical specification and reference implementation.

Category 3: High-Risk Consensus Changes

Soft forks that invalidate currently valid transaction types, restrict existing functionality, or could cause loss of funds for users with pre-existing transactions or scripts. BIP-110 falls into this category: it would invalidate transaction patterns that are currently valid — outputs whose scriptPubKey exceeds the new size caps (34 bytes for non-OP_RETURN outputs, 83 bytes for OP_RETURN), oversized data pushes and witness items, tapscripts that execute OP_IF/OP_NOTIF or contain OP_SUCCESS opcodes, witness stacks bearing a Taproot annex, spends of undefined witness or tapleaf versions, and oversized Taproot control blocks. The proposal exempts all pre-activation UTXOs and provides at least a two-week grace period between lock-in and activation, which greatly reduces the fund-loss risk. The

residual exposure is the narrow case of a pre-signed transaction spending an output created at or after the activation height (the pre-activation exemption snapshots at the activation height — block 965,664 on the mandatory path) whose only remaining spending path runs through a now-restricted tapleaf. That a high-risk consensus change must be reasoned about at this level of edge case is itself the point — it should rule out the scenario by design and by testing, not leave it to analysis. **Minimum review period: twenty-four months** from publication of a complete technical specification and reference implementation.

Category 4: Hard Forks

Any change that old nodes would reject. Hard forks carry the highest risk of permanent chain splits and should be avoided except in extraordinary cases, such as a critical security vulnerability discovered in the deployed protocol. Where they are pursued anyway, this framework proposes the following minimum standards: (1) a thirty-six-month review period from publication of a complete technical specification and reference implementation; (2) explicit replay protection or a published rationale for its absence; (3) demonstrated economic-node support, including affirmative commitments from major exchanges, custodians, and payment processors; and (4) a published chain-split contingency plan addressing user communication, exchange coordination, and the contractual question of which chain constitutes “Bitcoin” for outstanding obligations. The Bitcoin community’s strong norm against hard forks — reinforced by the block size wars — remains the most effective deterrent, but the lack of a written standard has historically left proponents free to define their own. The necessity of replay protection is itself the diagnostic of hard-fork status. A proposal whose safe deployment requires the ecosystem to change its transaction format to prevent cross-chain replay has, by that requirement, made the case for its hard-fork classification — and should be judged against the standards of this category rather than treated as a soft fork with a safety net bolted on. The standards that follow in §3.7 distinguish between scheduled and emergency hard forks, which have opposite governance properties despite sharing this category.

3.3 Code Audit Requirements

The activation client for any consensus change must meet the following code quality standards before activation signaling begins:

A. Diverse independent review. The reference implementation must be reviewed by at least three developers whose primary organizational affiliations differ from each other and from the proposal’s authors. “Organizational affiliation” for this purpose means current or recent (past 24 months) employer or material funding source. Reviewers must publicly disclose any prior collaboration with the proposal’s authors — co-authorship on prior BIPs or papers, joint employment, or shared funding. Disclosure does not disqualify; it makes the relationship legible. The standard is diverse independent perspectives, not pristine isolation, which is unachievable in a small developer community and weaponizable when

claimed. Reviewers should have demonstrated competence in Bitcoin protocol development, shown by prior contributions to Bitcoin Core, Bitcoin Knots, or another consensus-compatible implementation.

B. Test coverage. The reference implementation must include unit tests covering all new validation rules, integration tests demonstrating compatibility with existing valid transactions, and regression tests for all identified edge cases. Test results must be publicly reproducible.

C. Testnet deployment. The activation client must be deployed on a public Bitcoin test network (signet or testnet) for at least three months before mainnet activation signaling begins. That deployment must demonstrate successful activation, enforcement of new rules, and — critically — successful deactivation if the proposal includes a sunset clause.

D. Fuzzing and adversarial testing. The reference implementation should be subjected to automated fuzzing and adversarial testing to identify vulnerabilities that could be exploited during or after activation. This is particularly important for proposals that restrict transaction types, as attackers may attempt to craft transactions that exploit ambiguities in the new rules.

E. Reviewer comprehension. Every change to consensus-critical code must be reviewed by at least one named human reviewer who publicly attests that they understand the change and can defend its correctness against technical challenge. This requirement applies regardless of the code's origin — whether authored by humans, generated by AI tools, or adapted from prior proposals. AI-generated code in consensus-critical software is analogous to AI-generated legal filings: the tool can accelerate production, but the professional remains responsible for the output's correctness. The framework does not try to police the origin of code, which is undetectable on inspection; it requires named accountable reviewers to understand the code, which is testable in any review forum. Authors and reviewers are welcome to disclose AI involvement voluntarily for transparency, but the load-bearing requirement is reviewer comprehension, not origin disclosure.

These requirements are not written on a blank slate. Bitcoin Core's existing review discipline already enforces much of what this section formalizes: concept and approach review, then code-level acknowledgment, with merge reserved to a maintainer only once peer review has surfaced rough consensus, and a markedly higher bar applied to consensus-critical code. A change that has earned that scrutiny will satisfy most of these standards as a byproduct. The framework's contribution is to extend that discipline from the codebase to the activation campaign, where it has historically been absent: a client can be forked, promoted, and bundled into a release stream entirely outside the review process that governs the reference implementation — which is exactly what happened with BIP-110. Two structural facts bound what these requirements can achieve, and both should be stated plainly. The pool of reviewers competent to evaluate consensus-critical code is small and concentrated, which is why the standard is diverse independent perspective rather than the unreachable ideal of complete independence. And the network's reliance on a single dominant implementation is itself a systemic governance risk that no review standard removes; the diverse-review requirement

mitigates that concentration at the margin, by ensuring consensus-critical changes are not validated solely from within one project, but it does not pretend to solve a monoculture that is a property of the ecosystem rather than of any single proposal.

3.4 Activation Threshold Standards

The activation threshold for a consensus change should reflect both the risk of the change and the cost of a failed activation. These thresholds are proposed as minimum standards:

Activation Mechanism Design Space

Threshold selection is downstream of mechanism selection. The activation mechanisms employed in Bitcoin’s history form a small but instructive design space. Each represents a distinct trade-off between activation speed, fail-safe behavior, and the locus of signaling authority.

BIP-9 (Version Bits with Timeout and Delay). The first versionbits activation mechanism, introduced for the CSV soft fork (BIPs 68, 112, and 113) and used again for SegWit (BIP-141). It superseded the earlier IsSuperMajority mechanism, under which CLTV (BIP-65) and the soft forks preceding it had activated. Miners signal readiness via version-bit flags; the deployment locks in when the configured threshold of blocks within a 2,016-block window signal readiness, and activates one retarget period later. BIP-9 includes a timeout: if the threshold is not met within a defined window, the deployment expires. BIP-9’s principal weakness is that it grants miners effective veto power — a small mining pool coalition can block activation by refusing to signal, even where node operators, exchanges, and users overwhelmingly support the change. SegWit’s eight-month stall demonstrated this failure mode in practice.

BIP-91 (Reduced Threshold MASF). A direct response to SegWit’s stalled BIP-9 deployment. BIP-91 lowered the lock-in threshold to 80% and made signaling itself compulsory: BIP-91-enforcing miners would reject blocks that did not signal for SegWit. This created the coordination pressure that produced universal SegWit signaling in July 2017 and carried BIP-141 to lock-in in early August. BIP-91 demonstrated that compulsory signaling — not just threshold reduction — can break a deadlock, at the cost of accepting some chain-split risk during the coordination period.

BIP-148 (User Activated Soft Fork). A flag-day mechanism: BIP-148-enforcing nodes would begin rejecting non-SegWit-signaling blocks on August 1, 2017, regardless of miner readiness. BIP-148 transferred activation authority from miners to economic nodes. BIP-148 worked because overwhelming community support produced miner capitulation; the mechanism’s credible threat is widely credited with motivating the capitulation that produced BIP-91. UASF without that prior support yields the BIP-110 pattern — negligible voluntary

signaling and a forced-activation confrontation. The mechanism is structurally riskier than miner-activated alternatives — if miners refuse to comply, the BIP-148 chain splits from the legacy chain — and is appropriate only where the underlying proposal has overwhelming economic support.

Speedy Trial. The activation mechanism used for Taproot in 2021. A bounded BIP-9 deployment with a short signaling window (approximately three months) and a high threshold (90%). If the threshold is reached, activation occurs after a defined delay; if not, the proposal expires without activation, freeing the deployment slot for revision or alternative proposals. Speedy Trial trades the certainty of activation for the certainty of a defined timeline. It is the most cautious of the modern mechanisms and is the closest existing match to the standards proposed in this framework.

LOT=true / LOT=false. A parameter that arose in BIP-8 during Taproot’s activation debates. Lockin On Timeout (LOT) specifies whether a deployment that fails to achieve miner signaling within its window should nevertheless activate by user mandate at timeout. LOT=true converts a deployment into an effective UASF if miners do not cooperate; LOT=false makes it a clean Speedy Trial-style timeout. LOT=true is inappropriate for a first deployment attempt and contested even on repeat attempts. The LOT debate of early 2021 surfaced the central question of every activation mechanism: who holds final authority over consensus rule changes, and what happens when they disagree?

This framework treats activation mechanism selection as a deliberate design choice subject to the standards in this section. A proposal that meets the substantive standards but uses an inappropriate activation mechanism remains deficient. A proposal that uses an appropriate mechanism but fails the substantive standards is no more defensible. Mechanism and merit are both gating.

Miner-Activated Soft Fork (MASF)

Minimum threshold: 90% of hashrate measured over a defined signaling period of at least two weeks (2,016 blocks). This threshold matches Taproot’s successful activation and minimizes the risk of a chain split. Thresholds in the 80–90% band are risky but historically achievable, with most of the safety margin already gone. Thresholds below 80% should be considered presumptively dangerous. Thresholds below 60% are reckless and should be rejected regardless of the proposal’s merits.

The 90% floor is anchored, not arbitrary, but it is not the only defensible number. It is the lowest deployment threshold at which a modern soft fork has activated without a split (Taproot), and the threshold at which the model below drives six-block reorganization exposure below roughly two parts per million. BIP-91’s 80% is not a counterexample: it was a transient compulsory-signaling overlay coordinating BIP-141’s 95% deployment, not a rule-deployment threshold. A reader could argue for 95% on the same model; the framework

treats 90% as a floor, not a ceiling, and the band boundaries (80% and 60%) are anchored to the historical record and to the signal-to-enforcement gap rather than to any feature of the curve: 80% is the lowest signaling level a modern deployment has used for coordination (BIP-91), and below 60% the gap between signaling and enforcement can put the enforcing share under one-half, where divergence stops being transient exposure and becomes permanent. What distinguishes this from BIP-110's 55% is not that one number was chosen and the other was not — both were — but that 90% is defended against a published risk model and 55% was not.

User-Activated Soft Fork (UASF)

UASFs should be reserved for cases where a proposal has demonstrated overwhelming community support but miner signaling is blocked by a small number of mining pool operators acting against their users' interests. The original UASF rationale, articulated in BIP-148 and the contemporaneous bitcoin-dev mailing list discussions of early 2017, held that user-activated mechanisms transfer activation authority from miners to economic nodes when the two diverge. A UASF is an extraordinary measure. It should not be the default activation mechanism for any proposal.

A UASF should only proceed when: (1) the proposal has completed its full minimum review period; (2) the proposal has demonstrated broad support among economic nodes, exchanges, and major holders; (3) the UASF includes a defined activation date set at least six months in the future to allow time for preparation; and (4) the UASF proponents have published a detailed chain split contingency plan.

The 55% Problem

BIP-110's 55% threshold deserves specific discussion because it shows the danger of low thresholds. A 55% signaling threshold permits activation when the hashrate actually enforcing the new rules may not meaningfully exceed 55% — and, because signaling is not enforcement, may fall below half. If a share of post-activation hashrate produces blocks valid under the legacy rules but invalid under the new ones, activated nodes will reject those blocks, and the network's chain views diverge. Where enforcement holds a majority, the enforcing chain leads in expectation but suffers transient deficits with the probabilities modeled below. Where enforcement falls below half — the scenario a 55% signal cannot rule out — the enforcing chain falls permanently behind in cumulative proof of work, and enforcing nodes hold a minority chain that legacy nodes recognize as valid but do not follow, because it carries less cumulative work. The 95% threshold used by SegWit and the 90% threshold used by Taproot exist precisely to drive that probability toward zero. A 55% threshold makes a persistent split foreseeable.

A 55% threshold does not safely activate a soft fork. It produces the conditions historically associated with persistent minority chains: weeks or months of competing tips, ambiguous economic status for transactions confirmed on either side, and unresolved questions about

the meaning of “Bitcoin” in contracts and on exchanges. That P2SH used the same threshold in 2012 is not reassuring precedent—it is the opposite. P2SH’s 55% coinbase poll fell short of its threshold on the first attempt and the chain forked between enforcing and non-enforcing miners; the change activated only after it was re-cut to a hardcoded flag day (§2.1). The one time Bitcoin ran this threshold for a real consensus change it produced exactly the split this model predicts—and on a network worth a tiny fraction of today’s. The stakes have changed. The standards must change with them. This conclusion is consistent with the qualitative analysis in *BCAP* §3.5.2, which identifies a low percentage of Economic Node enforcement as producing the highest risk of chain split and which describes the mechanism—sub-overwhelming enforcement permitting unupgraded blocks to extend the chain—by which that risk is realized.

Quantifying the Risk

The danger of a 55% threshold can be quantified. Model post-activation hashrate as enforcing (share E) and non-enforcing (share $1 - E$). Treat block production as a Bernoulli process: each block is, with probability E , valid under the new rules and built on the enforcing chain; with probability $1 - E$, it violates the new rules and extends a competing non-enforcing chain. Enforcing nodes follow the most-work chain that is valid under the new rules; non-enforcing nodes follow the most-work chain regardless.

The difference in cumulative work between the chains is a random walk—the same walk Satoshi Nakamoto used in Section 11 of the Bitcoin whitepaper to bound an attacker’s probability of catching up from a deficit. Only the meaning of the shares changes. There the decisive share was an adversary’s hashrate racing the honest network in secret; here it is the share E openly enforcing the new rules, and the race is not adversarial—neither cohort withholds blocks, and both mine at their natural rate from a tie at activation. Encode the enforcing chain’s lead as a step that moves $+1$ with probability E and -1 with probability $1 - E$, so its expected drift is $2E - 1$ per block. For $E > \frac{1}{2}$ the probability that the non-enforcing chain ever reaches a lead of k blocks is exactly $(\frac{1-E}{E})^k$ (for $E < \frac{1}{2}$ the enforcing chain falls permanently behind; at exactly one-half the walk is recurrent—either way the probability is one). These are first-passage probabilities over an unbounded race—they bound the chance the non-enforcing chain *ever* reaches a k -block lead, so truncating to a finite activation window can only lower them. Within this model’s single-race idealization, the figures below are conservative upper bounds on the exposure in any given window; the optimism, as the simplifications below explain, lives in the idealization itself. A separate consideration runs the other way. In the adjacent double-spend setting, Rosenfeld and Grunspan and Pérez-Marco show that the realistic finite-confirmation distribution is negative binomial, and that the simple closed form can *understate* catch-up risk once an adversary’s parallel progress during the confirmation wait is counted. That correction was computed for the adversarial race, not this symmetric one, so it is a directional caution here rather than a recomputation of the figures. Neither refinement narrows the orders-of-magnitude gap between a low and a high

threshold — which is the point. Substituting for $k = 6$, the conventional confirmation depth for finality (many services credit sooner, which only raises the exposure):

- At $E = 0.55$, the probability of a six-block deficit on the enforcing chain during activation is approximately $(0.45/0.55)^6 \approx 0.30$ — roughly thirty percent.
- At $E = 0.90$, the same probability is approximately $(0.10/0.90)^6 \approx 1.9 \times 10^{-6}$.
- At $E = 0.95$, it is approximately $(0.05/0.95)^6 \approx 2.1 \times 10^{-8}$.

Figure 3.1 plots the relationship across the full range of enforcement shares, with the historical thresholds marked. The curve is the argument: between 55% and 95% the exposure does not taper — it collapses.

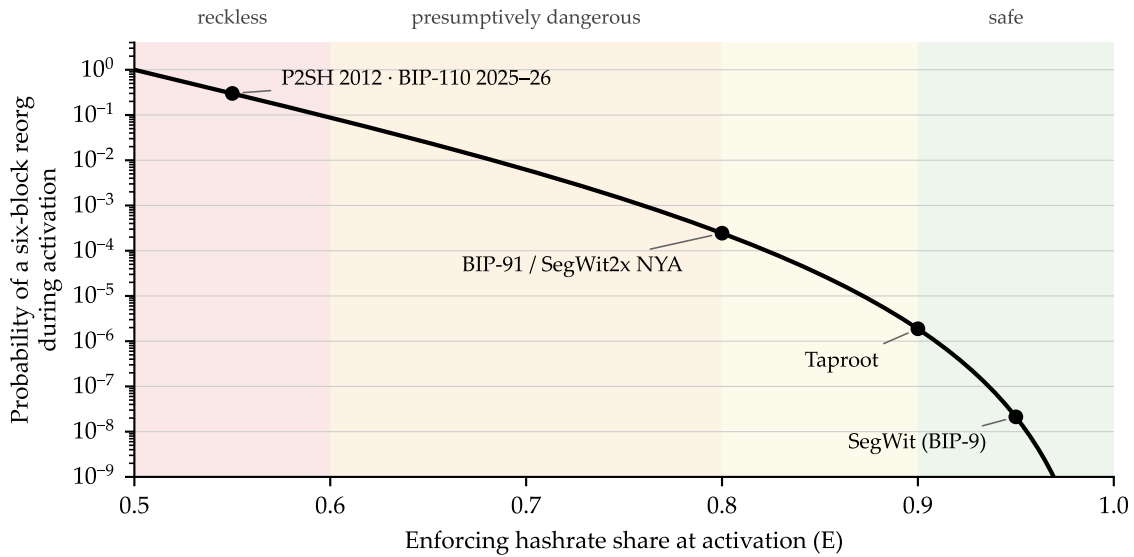


Figure 3.1: Probability that the non-enforcing chain leads the enforcing chain by six blocks at some point during activation, $P = ((1 - E)/E)^6$, log scale, as a function of enforcing hashrate share E . Historical thresholds are marked at their signaling values; the unlabeled band between 80% and 90% is §3.4's risky margin zone. The curve is a cliff, not a slope.

The perspective matters. Enforcing nodes never abandon their chain — they reject the competing blocks and simply fall behind. It is non-enforcing nodes, following the most-work chain regardless of the new rules, that reorganize onto the non-enforcing tip when it takes the lead — orphaning, from their view and from the view of every wallet, exchange, and service still running legacy rules, the transactions previously confirmed on the enforcing chain. Each six-block deficit is therefore an event in which deposits the legacy economy had credited as final lose their confirmations. At $E = 0.55$ the model gives roughly a thirty-percent chance that at least one such event occurs during activation; at $E = 0.90$ or above, such events are statistically nonexistent over the relevant time horizons. The 55%, 90%, and 95% thresholds

are not points on a linear spectrum of safety: they differ by five to seven orders of magnitude in expected reorganization exposure during activation.

This analysis simplifies several factors. It assumes non-enforcing miners produce blocks at the natural rate proportional to their hashrate; in practice, miners may defect to whichever chain becomes profitable, accelerating consolidation — and the more general lesson that miners can profitably depart from honest, natural-rate mining is the one Eyal and Sirer draw in *Majority Is Not Enough* (2014). It also assumes that every non-enforcing block contains a transaction that activated nodes would reject; for proposals like BIP-110 that target common transaction patterns, this assumption is largely realized, but for narrower changes the effective fork rate is lower. It also treats the race as a single episode from activation; in practice each orphaned non-enforcing block can start a fresh fork from the current tip, so exposure accumulates over the activation window — a correction that, like the others, runs against low thresholds. These simplifications affect the absolute magnitudes; none changes the qualitative conclusion that low-threshold activations are quantitatively distinct from high-threshold ones in their risk profile. And the corrections run one way: because a 55% signal cannot rule out enforcement below half, and because defecting miners migrate toward the higher-work chain, the realistic exposure is not a stable thirty-percent event but a drift toward capitulation or a persistent split — the thirty-percent figure is the optimistic, fixed-share corner of the model, not its ceiling.

The model runs on the static number by design, not oversight. A go/no-go standard has to run on inputs observable before the decision it gates: signaling share is measurable *ex ante* — anyone can count signaling blocks — while mid-fork price response, political pressure, and the timing of miner defections reveal themselves only after the network is committed. So the floor is computed on the worst-case snapshot of the one variable the network can verify in advance, the way a bridge is rated for the weight it must hold, not for the traffic that might reroute around it. The floor's height covers the proxy's slack as well: signaling only estimates enforcement, and a 90% signal leaves margin for the two to diverge where a 55% signal leaves none — slippage from 55% can land below one-half, where divergence is certain. Nor is the migration story a license for optimism: profit-driven hashrate movement is not guaranteed to run smooth or one way. When Bitcoin Cash split off in August 2017, its emergency difficulty adjustment produced months of hashrate oscillating between the two chains until a replacement algorithm shipped that November, and minority chains have persisted for years against the profit gradient — Ethereum Classic since 2016, Bitcoin Cash since 2017. If post-activation dynamics improve on the snapshot, the standard was conservative. If they do not, the snapshot was the only number anyone had at the moment of decision.

§3.4 in brief

The activation threshold is the share of mining power that must signal readiness for the new rules before they lock in. Higher threshold means less chance the network splits during activation.

- **90% or above** (SegWit, Taproot): splits are statistically nonexistent. Safe.
- **80–90%**: risky but historically achievable. Most of the margin of safety is gone.
- **60–80%**: presumptively dangerous. Splits become foreseeable.
- **Below 60%** (BIP-110's 55%): reckless. About a three-in-ten chance of a six-block reorg during activation if only 55% enforces — which a 55% signal does not assure. At 95%, the same chance is effectively zero — two parts in a hundred million.

These thresholds do not sit on a smooth curve. Between 55% and 95% the exposure shrinks by factors of millions — a collapse, not a discount (Figure 3.1). “55% is only 35 points below 90%” is the wrong way to think about it.

And the math runs on a share fixed at the pre-activation snapshot on purpose — signaling is the one proxy for it anyone can verify before the network commits. Everything after that is a forecast.

3.5 Chain Split Risk Assessment

The chain-split mechanisms this assessment must address are developed in detail in *BCAP* §3.5.2, including the bounty-claim scenario in which assets locked into scripts using newly proposed rules generate incentives for miners to mine an unupgraded block that voids the new-rule protection. Proposals that introduce new opcodes or new spending paths through OP_SUCCESS substitution warrant particular attention to that risk. The risk-assessment requirements that follow are the documentary correlates of that mechanism — questions a proponent must answer in writing before the community can rationally evaluate exposure.

Every consensus change proposal should include a formal chain split risk assessment addressing, at minimum:

A. Hashrate distribution analysis. What percentage of current hashrate is operated by pools or miners likely to adopt the change? What percentage is likely to reject it? Is there a credible path to the activation threshold, or is the proposal being pushed despite inadequate support?

B. Economic node analysis. Have major exchanges, payment processors, and infrastructure providers indicated support for the change? A consensus change that activates without exchange support creates immediate economic disruption, since users cannot deposit or withdraw funds until exchanges upgrade.

C. Replay protection. If the proposal would require explicit replay protection to prevent transactions from being valid on both chains in the event of a split, the proposal has crossed into hard-fork territory under Category 4 (see §3.7 on hard-fork subtypes) — reassess the classification before continuing the chain-split risk assessment. If the proposal is a soft fork that cannot produce a chain split absent miner defection, document the rationale for the absence of replay protection and address the expected impact on users in the residual miner-defection scenario.

D. Contingency plan. What happens if the activation fails? What happens if the activation succeeds but produces a persistent minority chain? Who is responsible for communicating the split to users, and how?

Contingency-plan contents

A chain-split contingency plan sufficient to satisfy criterion 19 (§5.1) should address, at minimum: (1) split-detection criteria and the trigger for invoking the plan; (2) the chain-naming and ticker convention to be applied — which chain is treated as “BTC”; (3) a user-communication plan specifying channels, timing, and explicit guidance on what holders should not do during the split; (4) an exchange and custodian coordination protocol covering deposit and withdrawal halts, confirmation-count adjustments, and crediting policy; (5) replay-risk disclosure and user-facing mitigation guidance; and (6) a wind-down or resolution path describing how the contingency state ends. The framework scores the existence and adequacy of such a plan, not its specific content, which is necessarily proposal-dependent.

3.6 Sunset and Reversibility Requirements

Proposals described as “temporary” must include a self-executing sunset clause. This means the consensus rules imposed by the proposal must automatically expire at a defined block height or timestamp without any further community action. The burden of continuation should fall on proponents of the change, not on opponents.

Specifically, a valid sunset clause must:

- A.** Define an exact block height or median time past (MTP) at which the new rules cease to be enforced.
- B.** Be implemented in the activation client such that nodes automatically revert to pre-activation consensus rules upon reaching the sunset trigger.
- C.** Be tested on testnet to confirm that deactivation works correctly and does not itself produce consensus failures.
- D.** Not require a later soft fork, hard fork, or software update to carry out deactivation.

A proposal that describes itself as temporary but requires active intervention to expire is not temporary. It is permanent with a stated aspiration.

3.7 Hard Fork Subtypes: Scheduled vs. Emergency

The framework’s general standards treat Category 4 (Hard Forks) as a single category requiring the highest scrutiny. In practice, hard forks fall into two extremes with opposite governance properties, and a single set of standards cannot accommodate both. The framework therefore distinguishes them: subtype 4a, the scheduled hard fork, and subtype 4b, the emergency hard fork. Each carries its own standards. The Category 4 text gestures at the emergency case by example (“a critical security vulnerability discovered in the deployed protocol”) but, applied uniformly, would impose a thirty-six-month review period on a vulnerability that cannot wait three years. The subtype distinction resolves that tension.

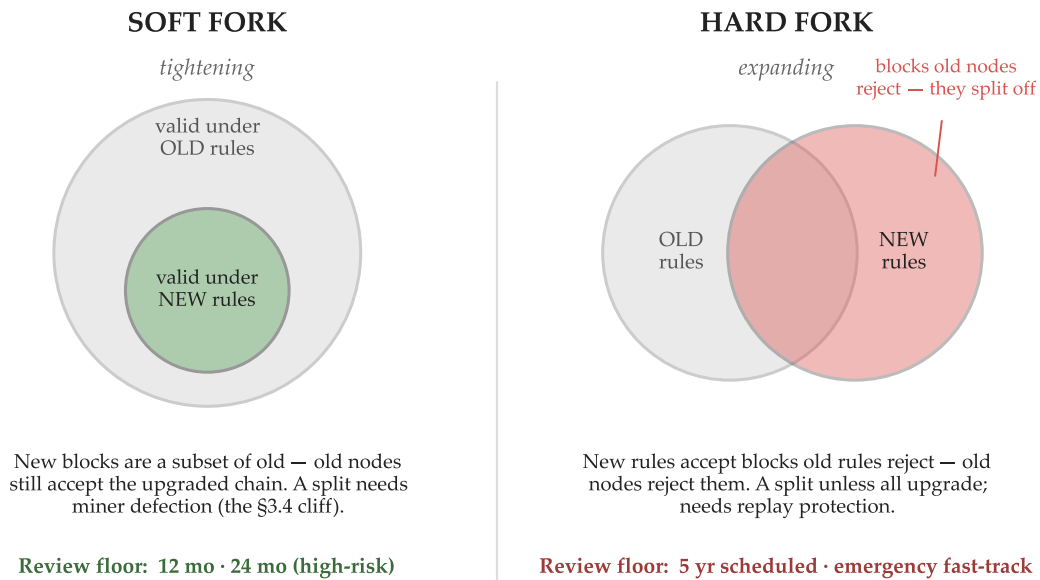


Figure 3.2: Soft fork versus hard fork, as a relationship between sets of valid blocks. A soft fork tightens the rules, so every block valid under the new rules is still valid under the old (a strict subset); un-upgraded nodes keep accepting the upgraded chain, and a split occurs only if too little hashrate enforces the change (§3.4). A hard fork expands the set of valid blocks, so the new rules accept blocks the old rules reject; un-upgraded nodes reject those blocks, producing a split unless every node upgrades — which is why replay protection matters. The minimum review-period floors scale with that risk.

4a — Scheduled Hard Fork

A scheduled hard fork has a deployment horizon measured in years from initial proposal to activation. The defining property is that time is the principal asset: organic node-upgrade cycles can do most of the coordination work if the deployment window is long enough.

Empirical analysis of historical node-upgrade rates (see §1.5) shows 95% node-operator adoption taking roughly two years, possibly three for the most recent releases; a five-year window is set deliberately above even that upper estimate, to leave margin for the slower upgrade dynamics a hard fork's higher stakes invite and to absorb the change without active coordination. A scheduled hard fork that takes advantage of this property can be deployed with much less coordination burden than an accelerated hard fork attempt, because the upgrade arrives organically through routine maintenance cycles.

The standards for subtype 4a follow Category 4 generally, with refinements:

A. Minimum review period: five years from publication of a complete technical specification and reference implementation. This exceeds the thirty-six-month floor that Category 4 sets generally (and Category 3's twenty-four-month floor); the extra review window reflects the deployment horizon's premium on getting the change right the first time.

B. Broad consensus required prior to code freeze , not merely prior to activation. The deployment window operates only as designed if the code shipped at the start of the window is the code that will be enforced at the end.

C. Explicit deployment timeline published in advance with each milestone identified: specification freeze, code freeze, infrastructure-test deployment, and mainnet activation block height.

D. Replay protection may be unnecessary if the upgrade window genuinely absorbs the change—that is, if the organic upgrade rate at the activation height is sufficient that no meaningful minority continues to enforce the prior rules. Replay protection becomes necessary only if the active-coordination component re-enters as a fallback.

4b — Emergency Hard Fork

An emergency hard fork is triggered by a critical vulnerability or imminent failure mode in the deployed protocol that cannot wait for a multi-year deployment window. Examples include the 2013 chain split that produced BIP-50, an inflation bug analogous to CVE-2018-17144 for which the soft-fork mitigation path is unavailable, or a cryptographic primitive becoming compromised in a way that affects existing UTXOs. The defining property is that time is the principal enemy: the alternative to a fast hard fork is running broken software.

The standards for subtype 4b are not lenient versions of Category 4's; they are calibrated to a different problem. Imposing the scheduled-HF standards on an emergency case would be actively dangerous; imposing the emergency standards on a scheduled case would discard the safety margin the long horizon exists to provide. The 4b standards are:

A. Documented evidence of the trigger. The proposal must identify the specific failure mode, the timeline within which the failure becomes operational, and the technical basis for the timeline. "Critical vulnerability" without published technical detail is not sufficient.

B. Compressed expert-led review. The review timeline is calibrated to the threat. The standards of §3.3 (Code Audit Requirements) apply — at least three reviewers from distinct organizational affiliations, named human comprehension attestation, testnet deployment — and the required duration is the minimum that lets these standards be met, typically weeks to months depending on the threat’s timeline.

C. Near-unanimous infrastructure coordination at the activation moment. Major exchanges, custodians, payment processors, and mining pools must commit to the activation timeline together. Soft commitments are insufficient; the coordination must be sufficient that a chain split does not produce a viable competing chain.

D. Replay protection where a meaningful minority will continue running the prior rules. In emergency-HF cases, the choice not to upgrade is itself often a principled position — an operator who disagrees that the trigger justifies the fork — and the prior-rules chain may persist with non-trivial economic activity. Replay protection in this scenario is not a category error; it is necessary user protection.

E. Transparent post-hoc disclosure of the decision process. The compressed timeline necessarily limits the breadth of pre-deployment consultation. The legitimacy of an emergency hard fork is established afterward by publishing what was decided, by whom, on what evidence, and with what dissents — making the process accountable in retrospect even where it could not be open in real time.

These standards assume a genuine emergency, and telling a real one from a manufactured one is itself part of the test. Three conditions separate them. The trigger must be independently reproducible: reviewers with no stake in the proposal must be able to confirm the failure mode from the published evidence rather than take the proponent’s word for it. The trigger must be a defect in deployed code — a way the protocol fails to do what it is meant to do — rather than a policy disagreement about how the protocol is being used. And the no-action path must be running broken software, not tolerating disfavored transactions. A use that one faction dislikes is not a vulnerability, and a proposal cannot purchase a compressed timeline by describing it as one. This is why a restriction proposal such as BIP-110 gets no review-period relief under the framework: its asserted trigger is the volume of inscription and data-carriage activity, which the consensus rules it proposes to change permit by design, so it is a Category 3 soft fork held to the full twenty-four-month floor (§3.2) and would not qualify for the emergency-hard-fork path even if it were re-pleaded as one.

The framework does not anticipate emergency hard forks being common: Bitcoin’s deployed protocol is exceptionally well-reviewed, and the post-2009 record contains no inflation-bug-level emergency that required this path. The March 2013 incident (BIP-50) was resolved in the emergency by a coordinated downgrade; the underlying incompatibility was retired months later on a scheduled flag day — a transition widely characterized as Bitcoin’s one de facto hard fork, and an uncontentious one. Recognizing subtype 4b acknowledges that some

governance needs invert under emergency conditions, not that emergency conditions suspend governance; the standards above keep the emergency case as legible and challengeable as the scheduled one.

Chapter 3 in brief

Chapter 3 sets seven gates a change should clear before anyone considers activating it. The bar rises with how much the change could break — so the first move is to gauge the risk, and the review period scales from there.

1. **A complete proposal** (§3.1). A problem statement backed by data, a full technical specification, a backward-compatibility analysis, the activation mechanism, a rollback path, and reference code with tests.
2. **Review time scaled to risk** (§3.2):
 - a soft fork that only *adds* rules (e.g., Taproot) — twelve months minimum;
 - a soft fork that *invalidates* transaction types in current use (e.g., BIP-110) — twenty-four months;
 - a hard fork — five years if scheduled, or a threat-calibrated fast track for a genuine emergency (§3.7).

The floors are sized to how long the network actually takes to upgrade, not picked for effect.

3. **Independent review** (§3.3). At least three developers from different organizations, with any prior collaboration disclosed — diverse perspectives, not pristine isolation.
4. **A safe activation threshold** (§3.4). Ninety-percent miner signaling at the floor; below eighty is dangerous, below sixty reckless.
5. **A chain-split risk assessment** (§3.5). Who holds the hashrate, whether economic nodes are on board, the rationale for any absent replay protection, and a written contingency plan.
6. **A self-executing sunset** (§3.6). Anything called “temporary” must expire on its own at a set block height — otherwise it is permanence with extra steps.
7. **A higher bar for hard forks** (§3.7). Scheduled ones get years of runway; an emergency must document the threat, compress review under named experts, and line up nearly every major operator before it ships.

None of this is enforced by code — Bitcoin has no enforcer. It is a benchmark: a proposal that clears the gates deserves serious engagement; one that does not should draw skepticism, whoever is championing it.

Chapter 4

Legal Analysis

The legal implications of Bitcoin consensus changes are largely unexplored. This is partly because Bitcoin’s decentralized nature complicates the application of traditional legal frameworks, and partly because no chain split has yet produced litigation with reported opinions. But the absence of precedent does not mean the absence of liability. The following analysis applies established tort and contract principles to the specific risks created by deficient consensus-change activation.

4.0 Purpose of the Legal Analysis

This section analyzes the legal-risk landscape that already exists, whether developers acknowledge it or not. It is not an invitation to litigation. It is a map of the terrain — drawn so participants, developers included, can navigate it with informed caution rather than unknowing exposure.

A developer who follows the Section 3 standards has a strong defense against every theory analyzed below. The legal analysis is best read in that light: as a description of what good-faith developers can demonstrate they have met, not as a catalog of theories under which they might be pursued. Section 3 is the safety standard; Section 4 is the case for why following the standard is itself the best protection against the liability theories that already exist in tort and contract law. The relationship runs in both directions. The scorecard of Section 5 turns those standards into a dated, public record; where a change is activated against a documented finding that it was unready, that record supplies the notice and foreseeability the theories below depend on. Compliance is the defense; the score is the exposure.

This section serves the network as a whole by making the legal landscape readable to every stakeholder — operators, exchanges, custodians, and developers — who must reckon with it without specialized training. Litigation against developers acting in good faith would harm the protocol and the community; the framework’s purpose is to make such litigation less likely by setting a documented standard of care that conduct can be measured against in advance.

How a readiness score becomes legal consequence

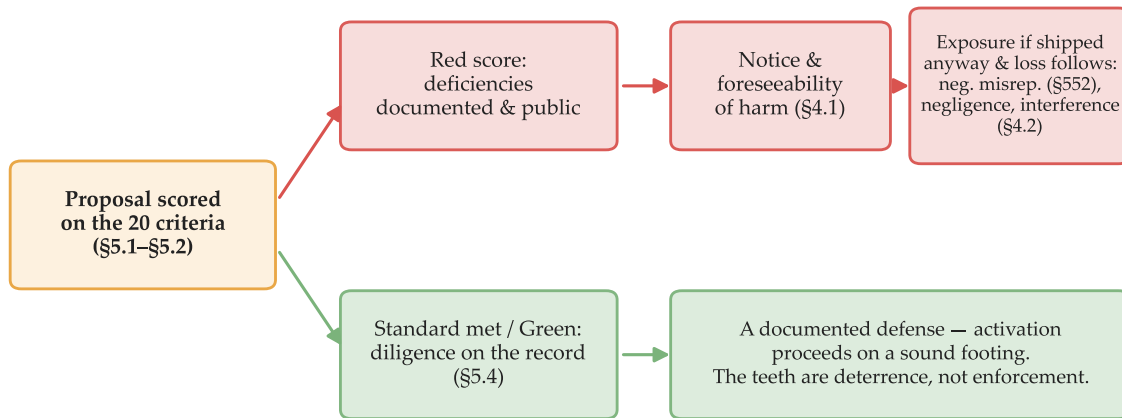


Figure 4.1: *How a readiness score becomes legal consequence. The framework’s force is deterrence, not enforcement (§4.0, §5.2, §6.5): a documented Red score is a dated, public record of a proposal’s deficiencies, and that record supplies the notice and foreseeability the negligence and negligent-misrepresentation theories of §4.1 depend on, should the change be activated anyway and cause loss. Following the standard produces the opposite record — documented diligence that is itself the defense. The figure traces the paper’s conditional argument; it asserts no liability outcome.*

4.1 Negligence

Tort liability for negligence requires a duty of care, a breach of that duty, causation, and damages. The first question is whether the developers of a consensus change activation client owe a duty of care to node operators and users who run their software.

Under traditional tort principles, a person who creates a dangerous instrumentality and places it into the stream of commerce owes a duty of care to foreseeable users. An activation client for a Bitcoin consensus change is software that, if defective, can cause direct financial harm to its users. The analogy to products liability is imperfect — most activation clients are distributed as free, open-source software — but open-source licenses do not categorically eliminate tort liability, particularly where the developer actively encourages adoption and knows that defects could cause financial loss.

The legal-exposure landscape

*When a negligently activated change splits the chain and people lose money,
more parties are exposed than most realize:*

Negligence / neg. misrep.	§4.1	Releasing activation code with known bugs and promoting its adoption; §552 routes around the economic-loss rule (Biakanja on contested facts).
Tortious interference	§4.2	A proponent with documented economic exposure who pushes a splitting change — especially if the promotion looks instrumental rather than principled.
Fiduciary duty	§4.3	Unsettled, but Tulip Trading (UK Court of Appeal, 2023) confirmed the question is justiciable (claim discontinued 2024); the developer's process determines it.
Pool operators	§4.4	Pools owe care over signaling, but broad-discretion pool terms blunt the contract/covenant route; exposure runs through what the pool tells its miners.
Regulatory & exchange	§4.5	Splits trigger income recognition on forked coins (Rev. Rul. 2019-24), custodial conflicts, and exchange listing calls — not hypothetical; this happened in 2017.
Following the Section 3 standards is a strong defense against every route. Compliance is the defense; the score is the exposure (§4.0).		

Figure 4.2: *The legal-exposure landscape this chapter maps. The five theories under which an inadequately reviewed activation can create exposure, each with its trigger—and the common defense that runs through all of them: following the Section 3 standards. Each route is conditional and developed, with authorities, in the section cited; the figure maps the terrain and asserts no liability outcome.*

Worked example: how the law below reaches a real person

Picture a node operator — say, a small exchange's infrastructure lead — who upgrades Bitcoin Knots through the routine version menu (§1.3), takes the activation client's stated "broad community support" at face value, and begins enforcing the new rules. Enforcement turns out to run far below the share the rules needed: her node follows what becomes the minority chain, a six-block reorganization unwinds a deposit her exchange had already credited as final, and a customer's funds vanish with it. Nothing here is exotic — it is the §3.4 cliff happening to a real person. The duty question is the one just above; what remains is the law that decides whether she actually has a claim against the people who shipped and promoted that client — the economic-loss rule and the negligent-misrepresentation route around it (Restatement § 552) — and the same law's answer: a developer who met the Section 3 standards has the documented defense (§4.0).

A second doctrinal obstacle is worth noting. California's economic loss rule generally bars recovery in tort for purely economic harm absent physical injury, property damage, or a special relationship giving rise to an independent duty. See *Aas v. Superior Court*, 24 Cal.4th 627 (2000); the rule's continued vitality — and the narrowness of its exceptions — was reaffirmed

in *Sheen v. Wells Fargo Bank, N.A.*, 12 Cal.5th 905 (2022). The rule is a real obstacle to a negligence theory against open-source developers whose users hold no contract with them and whose harm is financial rather than physical.

Two routes around the rule are available, and they are not equally strong. The contested route is the multi-factor test of *Biakanja v. Irving*, 49 Cal.2d 647 (1958), as applied to purely economic loss in *J'Aire Corp. v. Gregory*, 24 Cal.3d 799 (1979): a duty of care to non-contracting parties where the defendant's conduct was intended to affect the plaintiff, the harm was foreseeable, and the policy factors support liability. *Sheen* confirms that the *Biakanja* factors reach parties not in privity — the developer-and-user posture here. The honest difficulty is the transaction. *Southern California Gas Co. v. Superior Court (S. Cal. Gas Leak Cases)*, 7 Cal.5th 391 (2019), reaffirmed *J'Aire* but held that foreseeability alone does not set meaningful limits on liability for purely economic harm, and that the special relationship the duty requires is generally predicated on a transaction — and between a developer who publishes free software and an operator who downloads it there is no transaction at all: no contract, no payment, no privity, not even the indirect contractual chain *J'Aire* relied on, where a restaurant sued the contractor its landlord had hired. The route is therefore genuinely contested. The argument for it is that building, naming, and promoting a client *for* a defined set of node operators who download and rely on it for a single decision is nearer *J'Aire's* intended-beneficiary posture than the stranger-bystanders *SoCalGas* turned away, and that *Biakanja's* first factor — intent to affect the plaintiff — is met without a contract; the argument against it is that promotion aimed at a class is not a transaction with that class, and a court may treat the absence of any transaction as dispositive, as *SoCalGas* did. This route operates within California's general duty framework, not around it: the *Biakanja* factors are the economic-loss-specific form of the policy analysis *Rowland v. Christian*, 69 Cal.2d 108 (1968), governs across negligence, and *Brown v. USA Taekwondo*, 11 Cal.5th 204 (2021), reserves a threshold special-relationship inquiry for duties to control *third* parties — nonfeasance — which a developer's own release and promotion of defective code, being misfeasance, does not implicate. But the negligence theory does not rest on this route; it rests on the second, which does not need it.

That second route is negligent misrepresentation, and it carries the analysis. Where the developer makes representations about the safety or readiness of an activation client on which users foreseeably rely, it supplies a recognized cause of action without confronting the economic-loss rule at all. That theory rests on Restatement (Second) of Torts § 552 — carried forward, with refinements, in Restatement (Third) of Torts: Liability for Economic Harm § 5 (2020), though California continues to apply the § 552 formulation through *Bily v. Arthur Young & Co.*, 3 Cal.4th 370 (1992). (Where the representation crosses from negligent to knowing, the rule recedes further even between contracting parties: *Robinson Helicopter Co. v. Dana Corp.*, 34 Cal.4th 979 (2004), exempts affirmative misrepresentation, and *Rattagan v. Uber Technologies, Inc.*, 17 Cal.5th 1 (2024), extends the exemption to fraudulent concealment where the concealment is independent of the contract and exposes the plaintiff to a risk beyond the parties' reasonable contemplation at formation.) Both routes require facts beyond the bare release of buggy software; both are available on facts of the kind BIP-110 presents.

BIP-110's activation client shows the potential for negligence liability. The client was released with known bugs. Developers publicly identified defects that could cause users to fork themselves off the network. Despite these warnings, the client was distributed and its adoption was promoted on social media. The representation point is no longer hypothetical. The release notes that carried the activation rules into Bitcoin Knots' default stream in May 2026 assert that the upgrade already has broad community support and advise operators that declining or downgrading does not reject it—statements made to induce adoption, in a period when voluntary block-signaling for the change stood below one percent (§1.3, item D). A readiness representation of that kind, relied on by an operator who suffers loss, is the precise fact pattern the § 552 route addresses. *Bily*'s limits do not blunt the point. Section 552 reaches the limited class of persons for whose guidance the information is supplied, in a transaction the supplier intends it to influence. Release notes addressed to a node implementation's operators, urging the upgrade decision itself, define that class and that transaction with a precision the indeterminate audit-report audiences that concerned *Bily* never had. The contrary argument is that release notes broadcast to everyone who downloads widely distributed free software are the indeterminate, world-at-large audience § 552 refuses to reach. The better reading meets it: § 552 fixes the class by the transaction the information is supplied to influence, not by privity alone, and the operators deciding whether to run this consensus change are a purpose-bound class addressed in a known transaction—nearer the intended audience § 552 protects than the open-ended public whose merely foreseeable reliance *Bily* declined to reach. That the class is large does not by itself defeat the theory, though a court weighing the number and character of the audience could see it the other way. Section 552 also constrains the theory on the supplier's side. It reaches information supplied in the course of the supplier's business, profession, or employment, or in a transaction in which the supplier has a pecuniary interest; one who steps entirely outside any such role—publishing code gratuitously and walking away—falls outside it. The campaign's identifiable actors do not. Those who distribute the activation client or urge its adoption supply a readiness representation in the course of their own business, and a proponent with documented economic exposure to the outcome supplies it in a transaction in which it has a pecuniary interest as well; the wholly disinterested contributor, who supplies it on neither footing, is the one the theory correctly leaves alone (§6.8). Three further elements decide a § 552 case, and the actionable representation must be framed to meet them. The operative statement is not the bare claim of "broad community support"—which an operator could check against the visible signaling rate, and which a court might discount as opinion—but the representation of *readiness*: that the client was fit to run, conveyed by carrying the activation rules into the default release stream and advising that declining or downgrading "does not reject it." So framed, it is best read as a statement of present fact the operator cannot independently verify before running the code, which meets both the justifiable-reliance and the opinion objections; and the loss—orphaned blocks, an erased deposit—is the realization of exactly the unreadiness misrepresented, satisfying § 552's requirement that the loss flow from reliance on the false information rather than from a collateral fact. If a user had run this client and suffered a financial loss—for example, by mining blocks on a minority chain that

were subsequently orphaned — a negligence claim against the client’s developer would face challenging but not insurmountable hurdles.

The damages element deserves the same concreteness. A split of the kind §3.4 models produces losses that are specific, documentable, and conventional in kind: deposits credited against confirmations that a reorganization erases; coinbase revenue mined on a minority chain and orphaned with it; funds frozen where a pre-signed transaction’s only remaining spending path runs through a now-restricted tapleaf (§3.2); and the settlement ambiguity of §4.2, in which payment tendered on one chain is refused as payment on the other. Each category is the ordinary subject matter of commercial damages — lost funds, lost revenue, failed settlement — reached by an unusual mechanism, and each is computable from the chain data itself. In novel-technology cases the element that is hardest to prove is typically duty, not damages; the routes above address the harder element.

Causation, not damages, is where a defense will press first. The argument is that the split did not proximately cause the loss — that independent choices intervened: miners electing which chain to extend, users electing to run the client, exchanges electing how to credit and list. It fails on the same fact that makes the activation negligent. Under California law an intervening act supersedes only when it is so highly extraordinary as to be unforeseeable; where the likelihood that third parties will act in the very way that produces the harm is the hazard that made the conduct negligent, their so acting does not break the chain. Restatement (Second) of Torts § 449. A 55% activation is negligent precisely because it makes a split foreseeable — and the split is constituted by those operator, miner, and exchange responses, whose probability §3.4 computes in advance. The responses are the realization of the risk, not an intervening cause that severs it.

The strongest argument against liability is assumption of risk: users who run experimental software on a production network are arguably assuming the risk of loss. But assumption of risk is an affirmative defense, not a bar to the existence of a duty. And the defense is weaker when the software is presented alongside production releases in a platform’s version management system, as BIP-110 was on at least one node management platform.

A practical objection precedes all of these theories: the author of the BIP-110 client is pseudonymous, of unknown identity and location, and may well be judgment-proof. A suit naming that author alone could founder on identification, service of process, personal jurisdiction, and collectability before reaching the merits. The analysis here is therefore aimed less at the pseudonymous author than at the identifiable actors a consensus-change campaign necessarily involves — mining-pool operators who signal on behalf of connected miners, platforms that distribute the activation client, and public-figure promoters with documented economic exposure to the outcome. Pseudonymity is also not a permanent shield: identity can be sought through discovery directed at the platforms, pools, and intermediaries through which the code and the campaign travel. The framework’s documentary requirements exist partly to create the record by which those identifiable actors’ conduct can be evaluated.

Two defenses available to those identifiable actors should be named, for completeness and because each has limits. Platforms that distribute an activation client will invoke 47 U.S.C. § 230; its shield was built for publishing third-party content, and the fit is contested where the claim sounds in the platform's own product and presentation choices — the dropdown design, the absence of risk disclosure at the point of selection — rather than in the content of the code. See *Lemmon v. Snap, Inc.*, 995 F.3d 1085 (9th Cir. 2021) (product-design claims outside § 230). Developers will invoke the First Amendment status of code as speech, see *Junger v. Daley*, 209 F.3d 481 (6th Cir. 2000), and publishing code is protected in ways that operating an activation campaign is not — the same line this framework draws between a filed proposal and an ecosystem campaign (§1.0). Neither defense is trivial; neither is a categorical bar.

4.2 Tortious Interference

A consensus change activation that causes a chain split could give rise to claims of tortious interference with contractual relations or business expectancy. Consider the following scenario: a business accepts Bitcoin as payment under a contract that specifies payment in “Bitcoin.” A chain split occurs, and the payor delivers coins on the minority chain. The payee argues that “Bitcoin” means the majority chain. The resulting dispute was proximately caused by the chain split, which was proximately caused by the activation.

The question of which chain constitutes “Bitcoin” after a split has no settled legal answer. During the Bitcoin Cash fork, exchanges and contracts generally treated the chain with the most accumulated proof of work and the greatest economic activity as “Bitcoin.” But this convention is informal and could be challenged.

Proponents of consensus changes that carry a material risk of chain split should consider whether their actions could expose them to tortious interference claims. This matters most when the proponent is a company or public figure whose advocacy for the change is well-documented and whose economic interest in its success is apparent.

Two elements warrant emphasis. Tortious interference with contract requires intentional acts designed to disrupt performance, or undertaken with knowledge that disruption is substantially certain to result — not merely conduct that has the foreseeable effect of doing so. See *Quelimane Co. v. Stewart Title Guaranty Co.*, 19 Cal.4th 26 (1998); *Pacific Gas & Electric Co. v. Bear Stearns & Co.*, 50 Cal.3d 1118 (1990). Tortious interference with prospective economic advantage requires, in addition, an independently wrongful act — conduct unlawful for reasons other than the interference itself. See *Della Penna v. Toyota Motor Sales, U.S.A., Inc.*, 11 Cal.4th 376 (1995); *Korea Supply Co. v. Lockheed Martin Corp.*, 29 Cal.4th 1134 (2003). On the facts this paper develops, that wrongful act is identified, not assumed: the readiness misrepresentation analyzed in §4.1, actionable in its own right, supplies the independent wrongfulness the prospective-advantage prong requires. A developer who promotes an activation client out of ideological conviction usually will not meet either standard. The live claims are those in which the proponent of a consensus change has documented economic exposure to the

outcome—a financial position whose value depends on which chain prevails—and the proponent’s promotion can fairly be characterized as instrumental rather than principled. Such cases are not hypothetical. The framework’s documentary requirements—problem statement, backward compatibility analysis, contingency plan—are themselves designed to create the record by which such inquiries can be conducted.

4.3 Fiduciary Duties

Some legal scholars have argued that Bitcoin developers owe fiduciary duties to Bitcoin holders, analogous to the duties owed by corporate directors to shareholders. The most developed version of this argument is Angela Walch, “In Code(rs) We Trust: Software Developers as Fiduciaries in Public Blockchains,” in *Regulating Blockchain: Techno-Social and Legal Challenges* (Hacker, Lianos, Dimitropoulos & Eich eds., Oxford University Press 2019), which contends that core protocol developers exercise discretionary authority over property interests in a way that triggers fiduciary obligations under traditional principles. The premise that argument turns on—that the decentralization of protocol governance is in part a myth—is itself developed in Walch, “Deconstructing Decentralization: Exploring the Core Claim of Crypto Systems,” in *Cryptoassets: Legal, Regulatory, and Monetary Perspectives* (Brummer ed., Oxford University Press 2019). The position is contested. The leading counter-argument holds that core developers are not fiduciaries at all, because no participant—developers included—can unilaterally enforce a change on the network: the structural absence of centralized control makes the fiduciary frame ill-fitting. See Raina S. Haque, Rodrigo Seira, Brent A. Plummer & Nelson M. Rosario, “Blockchain Development and Fiduciary Duty,” 2 *Stan. J. Blockchain L. & Pol’y* 139 (2019). That objection has real force against the passive maintainer of a protocol who merely accepts or declines patches. It has much less force against the fact pattern this paper addresses: an actor who builds an activation client, names and promotes it, and campaigns for its adoption exercises far more control over the outcome than the passive maintenance the critics describe—which narrows, rather than defeats, the duty question for the active-deployment case.

The fiduciary question reached an appellate court in the United Kingdom’s Tulip Trading litigation, and the holding is narrower than the controversy around it suggests. In *Tulip Trading Ltd v van der Laan* [2023] EWCA Civ 83, the Court of Appeal of England and Wales did not adjudicate whether Bitcoin Core developers owe a fiduciary duty to holders of lost Bitcoin; it held only that the claim was sufficiently arguable to survive a strike-out application and should proceed to trial. The court allowed the appeal as a whole. The parallel common-law tortious duty-of-care claim, which the claimant had pleaded as arising only if a fiduciary duty existed, was sufficiently connected to the fiduciary issue that the court let it proceed to trial alongside it. Neither duty was decided on the merits; both were held to raise a serious issue to be tried. Birss LJ observed that “[i]f the decentralised governance of bitcoin really is a myth, then . . . there is much to be said for the submission that bitcoin developers, while acting as developers, owe fiduciary duties to the true owners of that property” (at [91]).

The trial never took place. Tulip Trading discontinued the claim in 2024, after the *Crypto Open Patent Alliance v Wright* litigation ([2024] EWHC 1198 (Ch)) rejected Dr. Wright’s claim to be Satoshi Nakamoto — the claim on which Tulip’s asserted ownership of the relevant bitcoin depended. The discontinuance leaves the Court of Appeal’s holding intact and the duty question untried. The case thus stands not for the proposition that developers owe such duties, but for the proposition that the question is justiciable on appropriate facts.

Those facts matter here. The fact pattern in *Tulip Trading* — developers’ alleged failure to issue a patch enabling recovery of stolen Bitcoin — remains structurally distinct from the chain-split scenario this paper addresses, where alleged harm would flow from the active deployment of defective consensus-altering code rather than from a failure to act. That distinction bears on how the duty question would resolve at trial, not on whether it is justiciable; and it does not foreclose the California analysis in §4.1, which rests on routes around the economic-loss rule — particularly *Biakanja* and Restatement § 552 — that turn on facts of intent, reliance, and representation that *Tulip Trading* did not engage.

This paper does not take a position on whether developers owe fiduciary duties in general. However, the analysis is relevant to consensus change governance for this reason: if a developer promotes a consensus change, distributes an activation client, and the change causes financial harm, whether the developer’s conduct breached a duty will be judged against the standard of care exercised in the process.

A developer who follows a rigorous governance framework — adequate review, thorough testing, conservative activation thresholds, and transparent communication of risks — has a strong defense against any claim of breach. A developer who releases a buggy client with a low activation threshold and no review period does not.

4.4 Mining and Node Operator Liability

Miners and node operators who adopt activation clients also face potential liability. A mining pool that signals for a consensus change bears some responsibility for the consequences of that signaling, particularly if the pool operator has not communicated the risks to the pool’s users (individual miners).

For solo miners — a category that includes the author — the analysis is simpler: a solo miner who runs an activation client is assuming the risk of their own operation. But a pool operator who signals on behalf of thousands of connected miners has a duty to those miners to exercise reasonable care in evaluating the consensus change. Mining pool operators who signal for poorly reviewed proposals with low activation thresholds are taking risks with other people’s hashrate and, by extension, other people’s money. The most likely cause of action sounds in contract or the implied covenant of good faith and fair dealing, with a candid qualification: most pool agreements disclaim liability and reserve broad operator discretion over signaling, so a claim premised on an express reasonable-care obligation will often founder on the contract’s own terms. The implied covenant fills part of that gap, but its reach

turns on the agreement's terms and is narrower than it first appears. Where the agreement is silent, the covenant governs the good-faith exercise of the pool's discretion, and signaling for a consensus change that exposes connected miners to orphaned blocks is the kind of decision it reaches. But where the agreement expressly confers broad signaling discretion — as pool agreements typically do — the covenant cannot impose a reasonable-care obligation the contract withheld (*Guz v. Bechtel National, Inc.*, 24 Cal.4th 317 (2000)), and conduct the contract expressly authorizes cannot itself breach it (*Carma Developers (Cal.), Inc. v. Marathon Development California, Inc.*, 2 Cal.4th 342 (1992)); it then reaches only an exercise of discretion so far outside the parties' reasonable expectations as to frustrate the agreement's central purpose. And where the agreement frames that discretion as sole or unfettered, even that residual narrows to the vanishing point: a court will not imply a covenant at odds with an express, consideration-backed grant of discretion unless a literal reading would render the bargain illusory (*Third Story Music, Inc. v. Waits*, 41 Cal.App.4th 798 (1995); *Wolf v. Walt Disney Pictures & Television*, 162 Cal.App.4th 1107 (2008)) — which a fee-paying pool agreement is not. In that common case the live exposure runs less through the silent agreement than through what a pool affirmatively represents to its connected miners about the change — a readiness representation of the kind §4.1 addresses. Two limits are worth naming: the governing law and forum of any such claim are fixed by the pool agreement's own choice-of-law and forum-selection clauses, which frequently point outside California, so the analysis here illustrates the kind of duty these agreements create rather than predicting the result under any particular one. Negligence remains available as a backstop, subject to the economic-loss-rule limitations discussed above.

4.5 Regulatory Consequences

Beyond direct liability, a chain split creates regulatory uncertainty. Tokens on the two chains are treated as separate assets for U.S. federal tax purposes: under Revenue Ruling 2019-24, a holder who receives new-chain coins following a hard fork recognizes ordinary income at receipt, measured by fair market value when the holder obtains dominion and control, and takes that value as basis in the new asset — the original coins' basis is unchanged. (The ruling's income-at-receipt model forgoes the alternative commentators had urged — allocating existing basis across the chains, the model used for securities spin-offs.) The taxable event is receipt of the new coins with dominion and control, not the fork itself; a holder who never receives or controls new-chain coins recognizes no income. Exchanges may have to support both chains or face claims from customers who hold tokens on the unsupported chain. Custodians may face conflicting obligations to clients. The framework brackets the asset-classification questions a split also raises — whether a forked token is a security, a commodity, or neither under federal law — as orthogonal to the split-liability analysis and governed by a literature of their own; the exposure mapped here attaches to the conduct of the activation, not to the legal character of the resulting asset. For Bitcoin and its forks the federal answer has since firmed in the direction the bracket assumes: the SEC and CFTC's

joint interpretive release of March 2026, 91 Fed. Reg. 13,714 (Mar. 23, 2026), treats Bitcoin (and Bitcoin Cash) as a digital commodity and a no-consideration distribution — a gift or airdrop for which the recipient provides no money, goods, or services — as outside the securities laws' investment-contract analysis (a category that, by its terms, reaches a forked-coin distribution received without consideration, though the release does not address forks specifically) — a development that reinforces rather than disturbs the bracket, since the exposure analyzed here turns on activation conduct, not asset character.

These regulatory consequences are not hypothetical — they happened during the Bitcoin Cash fork in 2017, and the tax questions alone took the IRS two years to resolve. A governance framework that minimizes the risk of chain splits also minimizes regulatory disruption. This is a feature, not a byproduct.

4.6 Comparative Note: Common-Law Jurisdictions and EU Software Liability

The legal analysis in §4.1 through §4.5 draws primarily on California state law and federal authority. The analytical method — negligence, tortious interference, fiduciary duty, contract, and regulatory exposure — translates across common-law jurisdictions with varying doctrinal specifics. This note maps the California analysis onto its closest cousins, identifies the most consequential current international development, and is explicit about civil-law limits.

A. United Kingdom. The duty-of-care analysis in §4.1 has a direct parallel in English common law. The leading modern duty-of-care case, *Caparo Industries plc v Dickman* (House of Lords, 1990), establishes a three-stage test: (1) foreseeability of harm; (2) proximity between the parties; and (3) whether imposing a duty is fair, just, and reasonable. *Caparo's* structural parallel to California's *Biakanja* multi-factor test makes the §4.1 analysis substantially portable for UK readers. Negligent misstatement claims under *Hedley Byrne & Co Ltd v Heller & Partners Ltd* (House of Lords, 1964) supply a UK analogue to the Restatement § 552 route in California, where the defendant assumed responsibility through representations on which the claimant reasonably relied. As §4.3 explains, the Court of Appeal in *Tulip Trading* allowed the tortious duty-of-care claim to proceed to trial as arguable rather than deciding it on the merits; the appeal turned on the connected fiduciary-duty issue, and the court did not resolve the *Caparo* duty question on the facts — nor, following the claim's discontinuance in 2024, has any English court since. One post-*Tulip* development bears noting: the Property (Digital Assets etc) Act 2025 (in force December 2025) confirms by statute that a thing is not prevented from being an object of personal property rights merely because it is neither a thing in possession nor a thing in action — placing the property status of crypto-tokens, the premise beneath *Tulip Trading's* duty claims, on statutory rather than purely common-law footing in England and Wales.

B. Commonwealth common-law jurisdictions. Singapore, Australia, Canada, New Zealand, and other Commonwealth jurisdictions derive their tort frameworks from English common law but have evolved distinct doctrines. Australia notably departed from *Caparo*'s three-stage test in *Sullivan v Moody* (High Court of Australia, 2001), adopting a framework that has moved away from the *Caparo* stages. Canada retained a modified version of the predecessor *Anns* (House of Lords, 1978) two-stage approach, refined through *Cooper v Hobart* (Supreme Court of Canada, 2001). The §4.1 analytical method translates as a roadmap, but jurisdiction-specific case law would govern any specific dispute; readers in these jurisdictions should consult counsel familiar with local doctrine.

C. European Union software liability. The most consequential current development in international software developer liability is the European Union's revised Product Liability Directive, Directive (EU) 2024/2853, adopted in 2024 to replace the 1985 directive. The 2024 revision explicitly extends strict product liability to software. Member States must transpose the new directive into national law within roughly two years of its entry into force. After transposition, a developer of defective software — potentially including a Bitcoin consensus-change activation client — may face strict liability for personal injury or property damage caused by the defect, subject to defenses available under the directive. The directive expressly excludes free and open-source software developed or supplied outside the course of a commercial activity (Art. 2(2); Recital 14). The live question is therefore not whether non-commercial open-source software is covered — by the directive's own terms it is excluded — but where the "commercial activity" boundary falls: software monetized through paid support, integrated into a commercial product, or supplied in exchange for personal data may re-enter scope. This paper takes no position on where a particular activation client sits on that line. The practical implication is that, after transposition, EU jurisdictions will present a very different liability landscape than the negligence-centered analysis in §4.1. Developers and pool operators subject to EU jurisdiction should consult counsel familiar with the directive's specifics in the relevant Member State.

D. Civil-law jurisdictions. Civil-law tort doctrine — including the foundational Swiss Code of Obligations Article 41 general clause and similar provisions in other civil-law systems — operates on different conceptual premises than common-law negligence. The framework's procedural standards in Section 3 and Section 5 translate (they evaluate procedural conduct, jurisdictionally agnostic); the legal-exposure analysis in Section 4 does not directly port and would require adaptation to local doctrine. This paper does not attempt that adaptation.

The framework's standards in Section 3 and Section 5 are jurisdictionally portable. They evaluate the procedural conduct of proposal development and activation, independent of any particular legal regime. The legal-exposure analysis in §4.1 through §4.5 is California-specific and should be understood as a methodological model. Liability exposure exists, in varying doctrinal forms, in every jurisdiction with developed tort, contract, and regulatory law. Practitioners outside the United States should treat Section 4 as a roadmap for the questions to ask local counsel, not as the answers themselves.

Chapter 4 in brief

When a negligently activated change causes a chain split and people lose money, more parties are legally exposed than most people realize:

- **Negligence (§4.1).** Releasing activation code with known bugs and promoting its adoption can produce negligence claims; negligent misrepresentation under Restatement § 552 is the cleaner route around the economic-loss rule, with the *Biakanja* duty available on contested facts.
- **Tortious interference (§4.2).** A proponent with documented economic exposure who pushes a change that splits chains can face live claims, especially if promotion looks instrumental rather than principled.
- **Fiduciary duty (§4.3).** Unsettled, but *Tulip Trading* (UK Court of Appeal, 2023) confirmed the question is justiciable; the claim was discontinued in 2024 without trial. The developer's process determines the answer.
- **Pool operators (§4.4).** Signaling for thousands of connected miners is a decision pools owe care over, but broad-discretion pool terms blunt the contract and implied-covenant routes; the live exposure runs through what the pool represents to its miners—a §4.1-style readiness representation.
- **Regulatory and exchange exposure (§4.5).** Splits trigger income recognition on forked coins (IRS Rev. Rul. 2019-24), custodial conflicts, exchange listing decisions. Not hypothetical—this already happened in 2017.

This is not a threat directed at developers. A developer who follows the Section 3 standards has a strong defense against every theory above. Section 3 is the safety standard; Section 4 is the case for why following it is the most effective protection. §4.6 maps the California analysis onto UK and Commonwealth common-law jurisdictions and flags the EU's 2024 Product Liability Directive as the most consequential current change in international software-developer liability.

Chapter 5

Proposed Standards

This chapter provides two evaluation tools that work at different speeds. §5.0 is a fast screening test: seven red flags that surface immediately on inspection of a proposal and indicate that the full evaluation is warranted. §5.1 is the full evaluation: twenty binary criteria synthesizing the framework in Sections 3 and 4, with the scoring rules in §5.2. A proposal that meets all the standards deserves serious community consideration. A proposal that fails to meet them should be treated with skepticism; one that proceeds to activation without meeting them is a candidate for the independent economic-node response set out in §5.2.

5.0 Red Flags: Is This Proposal Risky?

The following red flags surface from public information about a proposal — its activation parameters, its release timeline, its review record, and the conduct of its proponents. They are observable without running code or modeling outcomes. A proposal that trips two or more needs the full §5.1 evaluation before any operator runs the activation client. The flags are gateways to the criteria below, not scores in their own right.

1. **Activation threshold below 80%.** Sub-80% thresholds are presumptively dangerous; sub-60% is reckless. The gap in chain-split exposure between 55% and 95% is measured in factors of millions, not percentage points. See §3.4 and Figure 3.1.
2. **Roughly six weeks or less from initial proposal to activation client.** The framework's minimum review periods are twelve or twenty-four months by soft-fork risk category, and longer for hard forks (§3.2, §3.7). Review periods well short of the floor have failed before — see §1.3 (BIP-110) and §2.2 (block size wars).
3. **Activation client bundled into the default release stream of a node implementation.** Economic nodes get defaulted into the change through routine upgrades rather than opting in. The governance failure is structural even where the runtime prompt is real. See §1.3, item D.
4. **Reviewers all share an employer, funder, or recent prior collaboration** that isn't publicly disclosed. The standard is diverse perspectives with relationships made legible, not pristine isolation. See §3.3, item A.

5. **“Temporary” proposal with no self-executing sunset**, or a sunset whose deactivation has not been tested on testnet. “Temporary” without a sunset is permanence with extra steps; an untested sunset is a promise, not a guarantee. See §3.6.
6. **Proponents have documented economic exposure to the outcome**, where promotion can fairly be characterized as instrumental rather than principled. See §4.2.
7. **No published chain-split contingency plan**. If activation produces a persistent minority chain, who communicates the split to users, exchanges, and counterparties — and how? See §3.5, item D.

A proposal that trips one flag may still merit serious consideration; tripping two or more triggers the full §5.1 evaluation; tripping four or more is, on its face, disqualifying. The list is intentionally short. A proposal can be problematic for reasons not listed here, and a flag that is technically not tripped is not a clean bill of health. The framework’s full criteria determine readiness; the red flags determine only whether the framework needs to be run.

5.1 The Consensus Change Readiness Checklist

Scope note. The checklist evaluates scheduled consensus-change proposals. A genuine emergency hard fork (§3.7, subtype 4b) cannot, by construction, satisfy criteria 8, 15, or 16; it is evaluated directly against the subtype 4b standards — whose compressed timeline is paid for with the documentary, review, and coordination requirements set out there — not scored here as if those criteria could apply.

A. Proposal Quality

1. Does the proposal include a clear, empirically supported problem statement?
2. Does the proposal include a complete technical specification sufficient for independent implementation?
3. Does the proposal include a backward compatibility analysis identifying all affected transaction types, scripts, and use cases?
4. Does the proposal include a fully specified activation mechanism with defined thresholds, timelines, and failure modes?
5. Does the proposal include a rollback procedure? If described as temporary, does it include a self-executing sunset clause?

B. Code Quality

6. Has the reference implementation been reviewed by at least three developers with demonstrated Bitcoin protocol expertise, drawn from distinct organizational affiliations, with any prior collaboration with the proposal's authors publicly disclosed?
7. Does the reference implementation include comprehensive unit, integration, and regression tests?
8. Has the activation client been deployed on a public test network (signet or testnet) for at least three months, with successful deactivation tested if a sunset clause is included?
9. Has the code been subjected to fuzzing and adversarial testing?
10. Has every change to consensus-critical code been attested as understood by at least one named human reviewer who can defend its correctness?

C. Activation Safety

11. Where the deployment includes a miner-signaling threshold, is that threshold at or above 90%?
12. If a UASF mechanism is proposed, has the proposal completed its full minimum review period and demonstrated broad support among economic nodes?
13. Has a chain split risk assessment been completed and published?
14. Does the proposal include replay protection or, where the proposal is a soft fork that cannot produce a chain split absent miner defection, a documented rationale for its absence?
15. Does activation signaling begin only after the proposal's minimum review period (criterion 16) has elapsed, with enforcement scheduled no less than six months after publication of the final activation client?

D. Community Process

16. Has the proposal completed the minimum review period for its risk category (twelve months for moderate-risk soft forks, twenty-four for high-risk soft forks, and for hard forks the applicable §3.7 subtype floor)?
17. Has the proposal been discussed in public forums with participation from a diverse cross-section of the community (developers, miners, node operators, businesses, users)?
18. Have major exchanges and infrastructure providers been consulted regarding the proposal's impact on their operations?
19. Has the proposal's author published a chain split contingency plan?

20. Have the proposal's proponents published, or substantively responded to, a structured evaluation of the proposal against a published readiness standard — this framework or a comparable one — with the scoring and its evidentiary basis public? Engagement with *any* published, challengeable readiness standard — a rival rubric, or a published refutation of this framework — satisfies it, not engagement with this paper alone.

Scorecard worksheet

Mark each criterion “Met” or “Not”; tally at the bottom; classification bands in §5.2.

#	Criterion	Met	Not
<i>A. Proposal Quality</i>			
1	Clear, empirically supported problem statement.	☐	☐
2	Complete technical specification; independently implementable.	☐	☐
3	Backward-compatibility analysis covering all affected tx types, scripts, and use cases.	☐	☐
4	Fully specified activation mechanism (threshold, timeline, failure mode).	☐	☐
5	Rollback procedure; self-executing sunset if labeled “temporary.”	☐	☐
<i>B. Code Quality</i>			
6	≥ 3 expert reviewers from distinct organizations; prior collaboration disclosed.	☐	☐
7	Comprehensive unit, integration, and regression tests.	☐	☐
8	Test-network deployment (signet or testnet) ≥ 3 months; deactivation tested if sunset included.	☐	☐
9	Fuzzing and adversarial testing performed.	☐	☐
10	Named human reviewer attests comprehension of consensus-critical code.	☐	☐
<i>C. Activation Safety</i>			
11	Miner-signaling threshold (where used) ≥ 90%.	☐	☐
12	UASF (if used) has completed full review and broad economic-node support.	☐	☐
13	Chain-split risk assessment completed and published.	☐	☐
14	Replay protection, or documented rationale where soft fork cannot split absent defection.	☐	☐
15	Signaling only after the review floor elapses; enforcement ≥ 6 months after final client.	☐	☐
<i>D. Community Process</i>			
16	Minimum review period for risk category (12 / 24 months; §3.7 floors for hard forks).	☐	☐
17	Public discussion across diverse stakeholders.	☐	☐
18	Major exchanges and infrastructure providers consulted.	☐	☐
19	Chain-split contingency plan published by proposal author.	☐	☐
20	Structured evaluation against a published readiness standard, published or answered by proponents.	☐	☐

Total: _____ / _____ applicable

Classification (§5.2): 100% = Green; 75–99% = Yellow; 50–74% = Orange; < 50% = Red.

5.2 Scoring

Each applicable criterion gets a binary score: met or not met. Proposals are classified by the share of applicable criteria met:

- **100%: Green.** The proposal has met every applicable minimum standard and is ready for activation signaling. Green is deliberately demanding: a single applicable criterion scored not met drops a proposal to Yellow, because *ready* means every standard met, not most of them — the band is a gate, not a curve. Contestation is not itself disqualifying: the criteria that turn on opposition (12 and 17) and the sustained-opposition factor (§5.3) score opposition that has gone *substantively unaddressed*, not the bare existence of disagreement — so a contested change that has engaged and answered its critics on the record can still reach Green, while one that shipped over them cannot.
- **75–99%: Yellow.** The proposal has met most standards but has known gaps that should be addressed before activation.
- **50–74%: Orange.** The proposal has serious deficiencies and should not proceed to activation signaling until they are resolved.
- **Below 50%: Red.** The proposal fails to meet minimum standards for serious consideration. A Red classification indicates the proposal is not ready for activation; the responses available to stakeholders who reach that conclusion include: (a) public documentation of the deficiencies measured against this framework; (b) independent decisions by economic nodes — exchanges, custodians, payment processors, and major holders — each declining to recognize the activated chain as “Bitcoin” for purposes of contracts, deposits, and withdrawals; (c) running non-signaling, non-activation client software; and (d) where activation proceeds despite these objections, publication of a chain-split contingency plan to ensure user safety and minimize the economic damage of the resulting fracture.

For a proposal to which all twenty criteria apply, these bands correspond to the original cut points: 20, 15–19, 10–14, and below 10.

Not-applicable scoring. Two narrow classes of criterion are excluded from both numerator and denominator. Temporal: criteria 19 and 20 cannot fairly apply to a proposal that activated before any published readiness standard existed. Structural: criterion 12 applies only where a UASF mechanism is used, and criterion 11 applies only where the deployment includes a miner-signaling threshold (a pure flag-day deployment is evaluated under criterion 12 alone). No other not-applicable scoring is permitted. A criterion that is inconvenient is not inapplicable, and an evaluator who excludes one outside these two classes has stopped applying the framework.

A few criteria deliberately overlap. A rushed timeline can fail both criterion 16 (the review floor) and criterion 15 (the signaling sequence that depends on it); a missing plan can fail both criterion 13 (the risk assessment) and criterion 19 (the contingency plan it should contain). The overlap is intended, not a drafting accident: the framework scores documented process gaps, and one root cause that leaves several distinct gaps in the record is more telling than

one that leaves a single gap, not less. The overlap never counts a deficiency the record does not independently show.

For reference, BIP-110 scores 3/20 under this framework—all twenty criteria apply to it: credit for having a technical specification, a defined activation mechanism, and a self-executing sunset clause (though one whose deactivation was not demonstrably tested on testnet); no credit for adequate review period, code quality, independent review, activation safety, or community process. The bundling of RDTs into the default Bitcoin Knots release stream in May 2026 (§1.3, item D above) does not alter the binary count but intensifies the failure of criterion 12: where a UASF requires demonstrated broad economic-node support before proceeding, the operational reality after the default inversion is that the upgrade ladder defaults economic nodes into the activation, rather than nodes affirmatively opting in. The criterion's failure is now structural—non-endorsement has been built into the operator's burden—not merely a documentary gap. Taproot scores 17/17 on the criteria applicable to it: criterion 12 is structurally inapplicable to a MASF deployment, and criteria 19 and 20 postdate the activation, so all three drop from the denominator. Every criterion that could have been evaluated in real time was met or exceeded.

A word on the limits of the Red-band response. The independent economic-node response it describes is an option available to stakeholders, not a prediction about what they will do. History counsels modesty here. In prior contested forks—the Bitcoin Cash split, the near-miss of SegWit2x—exchanges and custodians did not coordinate on governance principle so much as follow liquidity: they listed what had a market, hedged by supporting more than one chain, and let trading resolve which carried the name. Economic nodes answer to their customers and their own legal exposure, not to a rubric, and the framework does not assume otherwise. It does not claim to manufacture a refusal the market is unwilling to make. Its claim is narrower, and it is the true one: for operators already inclined to refuse a proposal they judge unready, a published scorecard lowers the cost of acting on that judgment and supplies a record that makes the refusal defensible after the fact (Section 4). What the framework changes is the cost and the legibility of refusal, not the underlying willingness to refuse—and that is not a small thing. An economic-node response of this kind is not hypothetical: it ran once, in 2017, when the credible threat that nodes would enforce SegWit without miner cooperation (BIP-148) overrode a miner majority and forced the activation. The framework does not invent that lever. It hands the next refusal a published standard to point at rather than a first-principles argument to mount under fire.

One legal note on the economic-node response itself, because the framework's critics will raise it if its supporters do not. Concerted refusals to deal among competitors can draw scrutiny under Section 1 of the Sherman Act, 15 U.S.C. § 1, and a standard-setting process captured by economically interested members sits squarely within the antitrust laws' historical attention—in *Allied Tube & Conduit Corp. v. Indian Head, Inc.*, 486 U.S. 492 (1988), a manufacturer packed a private standards body to exclude a rival and was denied antitrust immunity. That capture is the structural opposite of what this framework does, which is publish non-binding criteria each operator applies unilaterally. The framework's design is

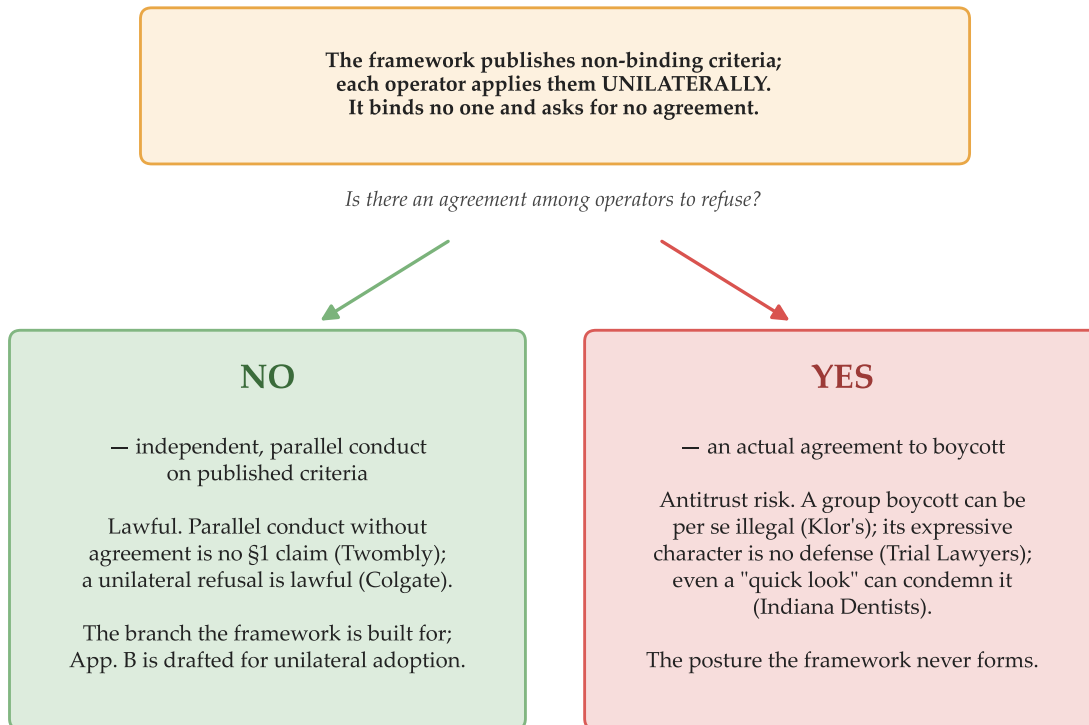


Figure 5.1: *The antitrust line at a glance. The framework supplies a shared evaluative input to a recognition decision each operator makes for itself; the dividing line is agreement. Independent parallel conduct on published criteria is lawful (*Twombly*; *Colgate*); only an actual agreement to refuse would raise the group-boycott concern (*Klor's*) — the posture this framework is designed never to form. The full analysis, with authorities, follows.*

the answer, and it is worth making explicit. The framework binds no one and asks for no agreement: it publishes criteria, and each operator applies them unilaterally to listing, custody, and risk decisions that exchanges already make, and must make, for every fork. Independent parallel decisions reached on published criteria are not a cartel — parallel conduct without agreement does not state a Sherman Act claim, see *Bell Atlantic Corp. v. Twombly*, 550 U.S. 544 (2007). An agreement among competitors to boycott would be a different thing. A literal agreement among exchanges to refuse a particular chain could draw *per se* scrutiny as a group boycott under *Klor's, Inc. v. Broadway-Hale Stores, Inc.*, 359 U.S. 207 (1959), and the published, expressive character of the criteria would not save it: *FTC v. Superior Court Trial Lawyers Ass'n*, 493 U.S. 411 (1990), held an economically motivated boycott *per se* unlawful despite its political and expressive content, which forecloses a "just sharing information" defense once an actual agreement exists. And an agreement that escaped *per se* treatment would not thereby be safe. *FTC v. Indiana Federation of Dentists*, 476 U.S. 447 (1986), condemns a concerted refusal on a "quick look" — proof of actual anticompetitive effect, with no full market-power showing (market power being "but a surrogate for detrimental effects"). That approach, like the others, presupposes the agreement this framework never forms. The framework's design is built to never reach that posture. It supplies a shared evidentiary input to decisions

each operator reaches independently, and two features keep it out of the *per se* box. The first is the absence of agreement: parallel conduct on common criteria is not concerted action (*Twombly*). The second is the market-power and exclusion screen of *Northwest Wholesale Stationers, Inc. v. Pacific Stationery & Printing Co.*, 472 U.S. 284 (1985), which distinguishes a cooperative arrangement with a plausible efficiency rationale from a naked boycott by parties with the power to exclude. Market power bears on that scrutiny only once concerted action is shown: a refusal to deal that no agreement underlies is independent action, lawful under *United States v. Colgate & Co.*, 250 U.S. 300 (1919). The participants' size therefore does not defeat the analysis; the absence of agreement carries it. A published common rubric also raises the *facilitating-practice* worry — that shared criteria ease tacit coordination among rivals — and the answer is the same in kind: a facilitating practice lowers the cost of converging on price or output and signals a firm's intentions to its competitors, whereas these criteria carry no price or output term, disclose no operator's hand, and feed each operator's own listing, custody, and crediting decisions without any exchange of assurances. Shared evaluative information that each firm applies unilaterally is the procompetitive kind, not the kind those cases condemn. Nor does the broader reach of Section 5 of the Federal Trade Commission Act, 15 U.S.C. § 45 — which can condemn an invitation to collude or a facilitating practice with no completed agreement at all — change the result: that reach stops where each firm's parallel conduct rests on an independent legitimate business justification, a ground on which *E.I. du Pont de Nemours & Co. v. FTC*, 729 F.2d 128 (2d Cir. 1984), set aside the Commission's facilitating-practice theory, and the ground that controls here, where declining to recognize an unready chain is independently rational for each operator at every fork. One feature still has to be confronted directly: the Red-band response does not merely share criteria, it recommends a convergent step — each economic node declining to recognize the activated chain — which is exclusion-shaped, the shape *Klor's* polices. The closest adverse authority is *Interstate Circuit, Inc. v. United States*, 306 U.S. 208 (1939), which inferred an agreement among competing distributors where an exhibitor invited each to adopt uniform terms and each adhered knowing the others had been invited and that the plan depended on common action. Three features separate this framework from that one. It issues no invitation to act in concert — it is published openly, conditions no operator's refusal on any other's, and asks none to act because the others will. It exchanges no assurances, and supplies only an evaluative input to a decision each operator must make for itself at every fork. And — the distinction *Interstate Circuit* turns on — declining to recognize a chain a published standard has flagged serves each operator's own interest whether or not any other joins, where the distributors' adherence served each one's interest only if all adhered together. A recommendation to decide independently, dependent on no other node's adherence and aligned with each node's own interest, is the absence of agreement, not its instrument. Nothing in this framework requires or contemplates an agreement, and adopted that way it stays on the lawful side of both lines. Operators adopting the framework should adopt it that way — as an input to independent judgment, documented as such — and the model language in Appendix B is drafted accordingly.

And the economic-node response is not the framework’s sharpest deterrent. The score is. A completed scorecard is a dated, public record that a proposal was measured against a known standard and found unready — and notice and foreseeability are precisely what several of the theories in Section 4 turn on. A change activated against a documented Red score is no longer an honest mistake: it proceeded after a published finding that it was reckless, available to everyone it would later harm. The routes around the economic-loss rule in §4.1 (*Biakanja*, Restatement § 552) turn on foreseeability and reliance; the tortious-interference theory in §4.2 turns on knowledge that disruption is substantially certain. A published Red score feeds each of them — and where the proponent has represented that the change is ready or has “broad community support,” the score is the proof of falsity the § 552 route turns on. The framework’s teeth, then, are not enforcement — Bitcoin has no enforcer — but deterrence: the documented score raises the legal and reputational cost of shipping a change the record already flagged. That is a narrower claim than “the framework stops bad proposals,” and a sturdier one — and unlike the economic-node response, it needs no one’s coordination. What it needs is a public, dated score, grounded in the underlying facts — the bugs, the 55% threshold, the sub-one-percent signaling — that documents the foreseeability those facts already carry; the score records the notice rather than manufacturing it, and its evidentiary weight grows as the framework is adopted (§6.5).

5.3 On the Measurability of the Criteria

The criteria were derived, not invented to fit.

Several criteria in §5.1 — particularly those concerning broad community support, sustained opposition, and adequate review — cannot be measured mechanically. There is no canonical metric for “broad support among economic nodes” or “sustained public objection” or “discussion in a diverse cross-section of the community.” Reasonable evaluators can score the same proposal differently.

This is intentional. The framework’s authority is its defensibility, not its precision. A score arrived at by counting blocks or polling exchanges would be precise but easily gamed; a score arrived at by transparent evaluation against published criteria, where reviewers must justify their scoring choices, is defensible against challenge in a way that algorithmic scoring cannot be. The fuzzy criteria force the conversation onto the right ground: what counts as evidence of broad support? What counts as sustained opposition? Who has objected, and have those objections been addressed? Those questions are themselves what the framework is asking the community to think about systematically.

For the fuzziest criteria, the heuristics below are illustrative anchors — not algorithmic rules. A proposal that satisfies the spirit of the criterion may not satisfy the literal heuristic and vice versa; the heuristic is a starting point for evaluators, not a substitute for evaluation.

“Broad support among economic nodes” (criterion 12, a precondition for UASF deployments) is reasonably evidenced by explicit public statements supporting the proposal from at least three of the top five exchanges by Bitcoin trading volume, plus comparable statements from at least two custodians serving institutional clients. Statements specifically endorsing the consensus change — not merely acknowledging it — count; silence does not count as support. Economic-node support resists direct measurement in the same way miner signaling overstates it: nodes do not broadcast their enforcement posture, and a public statement is only a proxy for an operating decision that is fully revealed at a split. The evaluator therefore works from the observable proxies — public endorsements, published node and listing policies, and, after any split, the chains that exchanges actually credit and quote — and treats their absence as it treats silence. This is the framework’s softest measurement point, and the posture of this section applies with full force: the score is defended in argument from the available evidence, not computed from a metric that does not exist.

“Sustained opposition” (a factor in the framework’s overall risk evaluation) is reasonably evidenced by at least two prior public objections from independent stakeholders that have not been substantively addressed in subsequent proposal updates. “Independent” for this purpose follows the diverse-affiliation standard of §3.3.A: objectors should not all share the same employer or material funding source.

“Adequate community discussion” (criterion 17) is reasonably evidenced by archived discussion in at least two distinct venues over the full minimum review period for the proposal’s risk category — typically the bitcoin-dev mailing list plus at least one of: a public forum thread, a recorded video discussion, or a working group with published minutes.

The remaining fuzzy elements in the scorecard’s earlier criteria are anchored as follows.

“Clear, empirically supported problem statement” (criterion 1) is reasonably evidenced by a one-sentence problem statement supported by at least one citation to quantitative data, measured user behavior, or on-chain observation. Aspirational framing (“Bitcoin should remain X”) is not a problem statement.

“Complete technical specification, independently implementable” (criterion 2) is reasonably evidenced where an independent developer could implement against the specification without consulting the author or reading the reference implementation. A reference implementation is a result of applying the specification, not a substitute for it.

“All affected transaction types, scripts, and use cases” (criterion 3) is reasonably evidenced by an explicit enumeration of every transaction type, opcode, and script pattern that becomes invalid or restricted under the change, together with a quantitative estimate of the affected on-chain population (UTXO count, value at risk, or comparable measure).

“Three expert reviewers from distinct organizational affiliations” (criterion 6) draws on two fuzzy elements. “Expert” is reasonably evidenced by at least one merged contribution to a consensus-compatible Bitcoin implementation, or by published technical analysis of Bitcoin consensus mechanics. “Distinct organizational affiliation” follows the standard of §3.3.A: current or recent (past 24 months) employer or material funding source, differing across reviewers and from the proposal’s authors.

“Comprehensive unit, integration, and regression tests” (criterion 7) is reasonably evidenced by test coverage of every new validation rule, every interaction with existing rules, and every identified edge case. Test results must be publicly reproducible by parties other than the implementation’s authors.

These heuristics are illustrative because the framework’s binary scoring is meant to be defensible against challenge, not bright-line. A reviewer who scores a criterion “met” using a different evidentiary basis than the heuristic above can defend that scoring by publishing the basis. A reviewer who scores a criterion “not met” can do the same. What the framework asks is not algorithmic agreement; it is that scoring choices be legible and challengeable. That is what makes the resulting evaluation defensible in any forum — technical, commercial, or legal — where the proposal’s readiness must be assessed.

On the closed loop. A sharper objection than score-by-score subjectivity is structural: one author selected the twenty criteria, chose which proposals to score, and assigned every mark — and the criteria were synthesized from the same activation history they grade, with the ninety-percent floor read off Taproot (§3.4) and the review and testnet floors off the same record. So Taproot at full marks and BIP-110 near zero are, the objection runs, foreordained by construction. The concession is owed and made: a rubric validated only against its author’s paradigm cases proves little. The answer is that the criteria were derived, not invented to fit. They cross-walk to independent prior sources — the thirteen evaluation questions and twelve indicators of *BCAP* (§1.5), Bitcoin Core’s own concept-then-code review discipline (§3.3), the IETF open-standards maturity gates that predate Bitcoin (§1.5), and the documented activation record — each of which reaches the same floors from outside this paper. That cross-walk is a derivation check — evidence the floors were not reverse-engineered from the verdicts — not an adversarial test, and should not be mistaken for one. The adversarial test is external: it is the dissent archive the conclusion commits to. The moment a published counter-scoring of Taproot or BIP-110 exists with its evidentiary basis — a defensible reading of the same record that places Taproot below seventeen or BIP-110 above three — the rubric is falsifiable on its own terms, and a closed loop cannot be falsified. SegWit2x (§5.4) is the early evidence the loop is not closed — a change the author’s own miner interest favors, scored Red against the same criteria.

Counter-scoring archive. To make that commitment more than a promise, the archive opens here, with the strongest counter-scoring the author can mount against his own results. Against Taproot’s seventeen, the sharpest attack is on criterion 13. The scorecard credits Taproot a published chain-split risk assessment, yet §3.5 defines

that criterion as a *formal* assessment — and this paper denies BIP-110 the same criterion for want of one. Taproot’s activation produced no single §3.5-form document either; its split-risk analysis was distributed across mailing-list threads, meeting logs, and the Speedy Trial proposal. A reviewer applying that strictness symmetrically could score criterion 13 *not met*, dropping Taproot to 16/17 — Yellow under the gate rule of §5.2, which makes any single not-met criterion band-moving. This is the sharpest of those attacks, and it is answered, not dismissed. Criterion 13 is the narrower of the readings the framework supports. Its components are scored elsewhere — replay protection at criterion 14, economic-node support at 18, the contingency plan at 19 — leaving 13 to ask only whether a split-risk assessment was completed and published at all. Taproot’s record satisfies that in substance: the roughly ninety-percent hashrate already supporting the change was public, and the LOT=true/false debate that dominated its activation was precisely a chain-split-scenario analysis. Criterion 13 is therefore *met* on the better reading and *contestable* on a stricter one; and even granting the stricter reading and a Yellow Taproot, the categorical gap this section turns on — every applicable criterion met, or all but one, against three of twenty — survives. A second band-moving attack targets criterion 18, on the same logic: the table credits Taproot major-exchange consultation, but that credit rests on the ecosystem’s demonstrated readiness rather than the formal, on-the-record consultation the New York Agreement produced for SegWit2x’s exchanges, so a strict reader could score 18 *not met* as well. It is answered as criterion 13 is — criterion 18 asks whether major exchanges and infrastructure providers were consulted on the proposal’s impact, which Taproot’s years of open review and the ecosystem’s broad, public readiness to support it satisfy in substance, and, granted or not, it moves Taproot only to Yellow, the categorical gap again surviving. Stack both strict readings — criteria 13 and 18 each scored not met — and Taproot still lands at fifteen of seventeen, comfortably Yellow; the gap this section turns on is between bands, not between marks, and the most hostile reading available leaves a Yellow Taproot against a Red BIP-110. A third, weaker attack turns the paper’s signaling-is-not-enforcement argument back on criterion 11: ninety-percent *signaling* no more guarantees ninety-percent enforcement than fifty-five guarantees fifty-five. That one fails outright — criterion 11 asks whether the signaling threshold is at or above ninety percent, which Taproot’s 1,815 of 2,016 blocks meets, and the signaling-to-enforcement gap decisive at fifty-five percent is immaterial at ninety, where the residual six-block-reorg exposure is on the order of two parts per million. Against BIP-110’s three, the most defensible move is to claim criterion 13 in the other direction: the specification describes the LOT=true window in which enforcing nodes reject non-signaling blocks, which a reader might take for a chain-split risk assessment. It fails too — describing an enforcement mechanism is not assessing the split risk it creates, and a release asserting “broad community support” while voluntary signaling sat below one percent is the opposite of a published assessment. These are real attempts, recorded so a better one can be built on them; the loop is closed only if no such attempt is ever made — this is the first.

5.4 Worked Examples: Taproot, BIP-110, SegWit2x, and a Neutrality Test

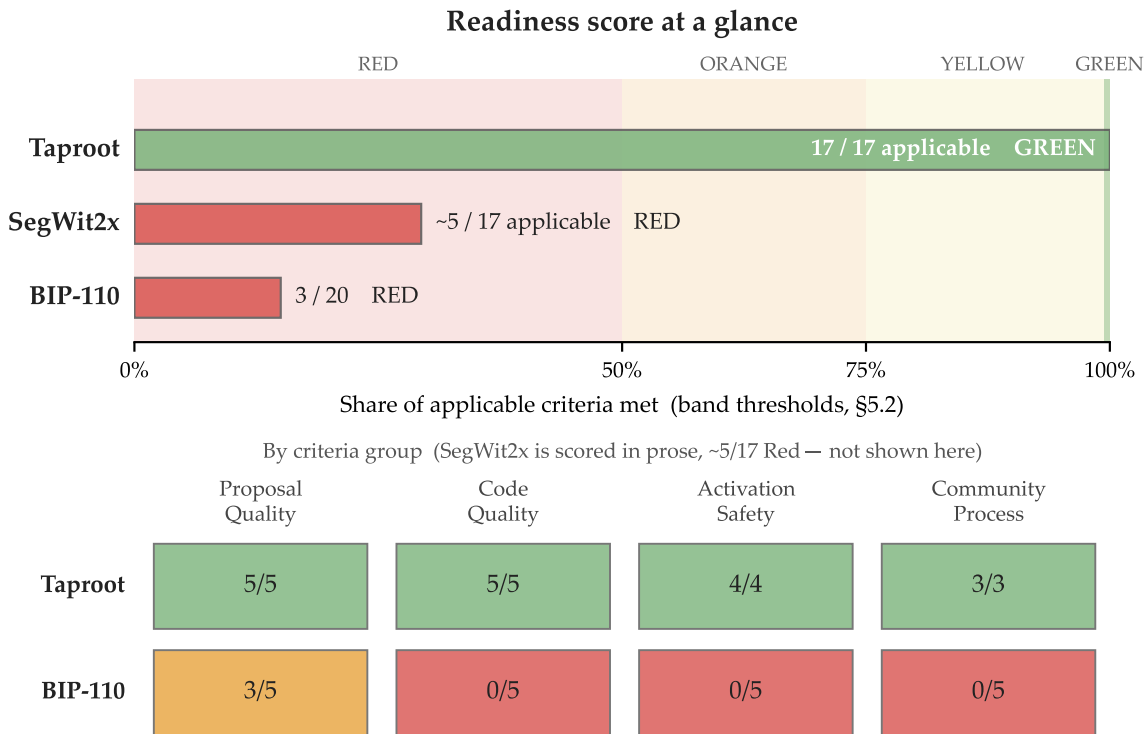


Figure 5.2: The worked examples at a glance. Each proposal is placed by the share of applicable criteria it meets, against the band thresholds of §5.2; the grid below breaks the two table-scored proposals into the four criteria groups. Taproot meets every applicable criterion (criterion 12 is structural and criteria 19–20 postdate its activation, so the three fall from the denominator); BIP-110’s three met criteria sit entirely within Proposal Quality; SegWit2x is a hard fork scored in prose at roughly five of seventeen. The criterion-by-criterion detail follows in the table.

The twenty criteria yield concrete classifications when applied to specific proposals. This section walks the criteria through three proposals: Taproot, which activated in November 2021 after years of review; BIP-110, whose late-2025 activation-client release attracted voluntary block-signaling of well under one percent; and SegWit2x, the 2017 block-size hard fork a coalition of miners and exchanges pushed and then withdrew. The first two are scored criterion by criterion in the table below; the third, and a fourth hypothetical example that follows it, are examined in prose, because together they test whether the framework measures process or just the author’s preference. The resulting scores — Taproot 17/17 on applicable criteria, BIP-110 3/20, SegWit2x roughly five of seventeen applicable — reflect categorically different framework readiness.

The scoring below applies the heuristic anchors of §5.3 and follows the framework’s commitment that scoring choices be “legible and challengeable.” A reader who reaches different

conclusions on the fuzzier criteria is invited to publish the alternative scoring with the evidentiary basis. The walkthrough’s purpose is to show the framework in operation, not to fix the score against further argument.

#	Criterion	Taproot	BIP-110
<i>A. Proposal Quality</i>			
1	Empirically supported problem statement	✓ Schnorr efficiency, script-path privacy, and key aggregation are measurable improvements	× “Spam uses” framed without quantitative measurement
2	Complete technical specification	✓ BIPs 340/341/342 permit independent implementation	✓ RDTs specification permits implementation
3	Backward compatibility analysis	✓ Strict soft fork; OP_SUCCESS slot reservation preserves future flexibility	× No enumeration of affected transaction types or use cases
4	Fully specified activation mechanism	✓ Speedy Trial: 90% threshold, ~3-month signaling window, defined timeout	✓ 55% UASF + LOT=true mandatory lock-in (Aug 2026); sunset ~1 year after activation
5	Rollback / sunset (if temporary)	✓ Permanent change; soft-fork status preserves future-reversal possibility per §3.1.E	✓ Self-executing sunset at defined block height
<i>B. Code Quality</i>			
6	≥ 3 expert reviewers, distinct organizations	✓ Multi-year review across Bitcoin Core, academic cryptography, and other consensus-compatible implementations	× Customary review burden not met; reviewer affiliations not documented (§1.3)
7	Comprehensive unit, integration, regression tests	✓ Full test coverage merged into Bitcoin Core’s test suite	× Client released with known bugs per §1.3
8	Test network (signet/testnet) ≥ 3 months; deactivation tested if sunset	✓ Multi-year signet and testnet deployment	× Six-week proposal-to-client timeline; sunset deactivation not demonstrably tested
9	Fuzzing and adversarial testing	✓ Bitcoin Core fuzzing infrastructure exercised on Taproot consensus code	× Fuzzing not documented
10	Named reviewer comprehension attestation	✓ Multiple named reviewers attested comprehension across years of public review	× No comparable attestation documented
<i>C. Activation Safety</i>			
11	MASF threshold ≥ 90%	✓ Speedy Trial threshold of 90% (1,815 of 2,016 blocks)	× 55% UASF threshold; sub-60% is reckless under §3.4

#	Criterion	Taproot	BIP-110
12	UASF: full review + broad economic-node support	n/a — Taproot deployed as a MASF; no UASF mechanism to evaluate (excluded from the denominator)	× Six-week review; < 1% voluntary signaling
13	Chain-split risk assessment	✓ Split-risk analysis in the public activation record (hashrate; the LOT split-scenario debate); defended in §5.3	× No formal chain-split risk assessment published
14	Replay protection or documented rationale	✓ Soft-fork status documented as the rationale for absence	× 55% threshold makes chain split foreseeable; no rationale documented
15	Signaling after review floor; enforcement ≥ 6 months from final client	✓ Signaling opened only after the twelve-month floor had run (complete specification and reference implementation public by early 2020); activation followed the final client (Core 0.21.1, May 2021) by six months	× Signaling opened immediately upon client release, six weeks into a twenty-four-month review floor
<i>D. Community Process</i>			
16	Minimum review period for risk category	✓ Approximately four years of review; far exceeds 12-month moderate-risk floor	× Six weeks vs. 24-month floor for high-risk consensus change
17	Public discussion, diverse stakeholders	✓ Sustained bitcoin-dev, conference, podcast, and working-group discussion	× Public discussion concentrated and brief
18	Major exchanges and infrastructure consulted	✓ Broad ecosystem readiness (exchange-operations impact minimal for a soft fork); defended in §5.3	× Exchanges not formally consulted; no public endorsements documented
19	Chain-split contingency plan published	n/a — contingency-plan format proposed by this framework postdates the activation (excluded from the denominator)	× No contingency plan published
20	Structured evaluation against a published readiness standard	n/a — no published readiness standard existed at activation (excluded from the denominator)	× No structured evaluation against a published readiness standard published or answered by proponents

Score: Taproot 17/17 applicable (Green). BIP-110 3/20 (Red). SegWit2x $\approx$ 5/17 applicable (Red; scored in prose below).

Taproot’s three excluded criteria reflect inapplicability, not defects: criterion 12 is structural (Taproot deployed as a MASF), and criteria 19 and 20 are temporal (no published readiness

standard existed in 2021). On the seventeen criteria applicable to it, Taproot scores 17/17 — Green — and frequently far above the framework’s floors: review period (years against a twelve-month minimum), reviewer breadth (a multi-organization community against a three-reviewer minimum), and testnet exposure (multi-year deployment against a three-month minimum).

BIP-110’s three “Met” results (criterion 2, technical specification; criterion 4, activation mechanism; criterion 5, self-executing sunset) sit entirely within the Proposal Quality category and reflect nothing more than the existence of a proposal document, a defined activation procedure, and a sunset block height. Every criterion in Code Quality, Activation Safety, and Community Process is not met, as are the remaining two Proposal Quality criteria (problem-statement empirical grounding and backward-compatibility analysis). The 3/20 score places the proposal in the Red band of §5.2, with the responses listed there available to stakeholders who reach that conclusion: public documentation of the deficiencies, independent decisions by economic nodes to refuse to recognize the activated chain as “Bitcoin” for contracts and deposits, running non-signaling node software, and publishing a contingency plan if activation proceeds anyway.

The gap between the two results — every applicable criterion met against three of twenty — reflects the framework’s substantive distinction: a proposal that received the rigor consensus-critical software requires versus a proposal that did not. The same framework, applied to the same kinds of evidence, produces categorically different results. A future BIP-110 variant that fixed even half the failed criteria would score differently — and the change would be visible in the record. The framework is a way of making such differences plain to every stakeholder in the network.

A third example completes the picture, and it is the one that matters most for the framework’s claim to neutrality. SegWit2x — the 2017 proposal to double the block size by hard fork, agreed at the New York meeting — was not a restriction advanced by inscription critics. It was the change miners and major exchanges pushed hardest: more block space, more fee-paying transactions, backed by a coalition that included Bitmain, Coinbase, Xapo, and BitPay, and signaled by over eighty percent of hashrate at its peak. It addressed a real, empirically grounded problem — full blocks and rising fees through 2017 — with an implementable specification and a defined activation height, and on Proposal Quality it scores well. Then it craters. Code Quality: zero — the btc1 client drew no independent review from the broader development community and no extended testnet exposure. Activation Safety: near zero — an eighty-percent threshold below the ninety-percent floor, no published chain-split risk assessment, and, decisively, no replay protection, the one feature whose absence would have had holders accidentally spending on both chains. Community Process: near zero — a six-month timeline against the five-year review floor a scheduled hard fork carries under §3.7 (and far inside even the thirty-six-month minimum Category 4 sets generally), and opposition broad enough that the proposal was withdrawn for lack of consensus. SegWit2x lands in the Red band — roughly five of its seventeen applicable criteria, concentrated in Proposal Quality, where it scores well, plus the exchange-consultation criterion (18) the New York Agreement

plainly met, with Code Quality, Activation Safety, and the rest of Community Process at or near zero — beside BIP-110. (Criterion 12 is structurally inapplicable to a miner-coordinated hard fork, and criteria 19 and 20 postdate the proposal.)

That result is the point of including it. The author of this framework is a miner; both inscriptions and bigger blocks put fees in his pocket, and his economic interest runs toward more demand for block space, not less. BIP-110 would restrict that demand and scores Red. SegWit2x would have expanded it and scores Red too. The framework flags the change that would have cost the author revenue and the change that would have paid him, for the same reasons in both: an activation below the safe threshold, no safety engineering, a rushed timeline, and a consensus that was never there. A scorecard that does not care whether a change serves the person holding it is the only kind worth holding — and that indifference, not the BIP-110 result, is the case for the framework. SegWit2x makes the point but does not make it cleanly. Its Red score is over-determined: it was a hard fork, and the framework flags any rushed hard fork on structural grounds — the sub-ninety-percent threshold, the absent replay protection, the timeline far inside even the five-year floor a scheduled hard fork carries (§3.7) — independent of whether the change served the author's interest at all. It was also the most-disfavored episode in the development tradition this paper draws on, which makes scoring it Red the easiest neutrality test, not the hardest. The sharper test is a change the author's own interest *wants*, flagged anyway, and a hypothetical isolates the author-interest variable.

Consider, then, a change from the camp opposite BIP-110's: a hypothetical *soft fork* that adds a new opt-in spend path — a new scripting capability of the kind Taproot reserved its `OP_SUCCESS` opcodes to allow, activating as a soft fork because un-upgraded nodes accept the new path as they always have — that draws a popular, fee-heavy new data-heavy use onto the existing block space, pushed by a coalition whose fee exposure runs toward greater block-space demand, the same direction the author's mining revenue runs. (Such a change adds a capability within the existing block budget; it neither raises the block's weight limit nor cheapens its data accounting, either of which would make it a hard fork, not a soft one — its draw is the new capability it adds, not any cheaper byte.) Suppose it arrives the way a rushed change arrives: an empirically thin problem statement (the new use is asserted to be wanted, not measured), a reference client released without independent review from outside the proposing coalition, a sub-ninety-percent activation threshold chosen to clear a contested change quickly, and no published chain-split contingency plan. The merits of the new capability are not the question — the framework scores process, not substantive desirability, and a careful version of this same change could score Green. This rushed version does not. It fails criterion 1 (no measured problem), the Code Quality criteria (no independent review, no testnet floor), criterion 11 (threshold below the ninety-percent floor), and the Community Process criteria (no minimum review period, no contingency plan). It lands in the Red band on the identical grounds that put BIP-110 there, and it lands there despite serving the author's fee interest and despite being precisely the kind of change BIP-110's proponents oppose.

This is the cleaner demonstration. BIP-110 restricts block-space demand and scores Red; SegWit2x and this hypothetical would expand it and score Red too — but the hypothetical, being itself a soft fork, removes the hard-fork structure that over-determined the SegWit2x result and removes the camp alignment that made it the easy case. What remains is a change the author’s own economic interest favors, opposed by the side BIP-110 sits on, and flagged Red on pure process. A framework that reds the change its author would profit from, on the same criteria it reds the change he would lose by, is measuring readiness and nothing else. That is a claim about internal consistency — the same facts score the same way whoever profits — not yet a proof of external neutrality, which only an adversarial counter-scoring by other hands can supply; the dissent archive (§5.3) is where that further test is run.

Chapter 5 in brief

Chapter 5 turns the framework into two tools.

- **A quick test** (§5.0). Seven red flags you can spot on sight — a threshold below 80%, a roughly six-week timeline, a client bundled into a default release stream, undisclosed reviewer ties, an untested sunset, conflicted promotion, and no contingency plan. Trip two or more and the proposal needs the full evaluation.
- **A readiness checklist** (§5.1). Twenty yes-or-no criteria in four groups — Proposal Quality (1–5), Code Quality (6–10), Activation Safety (11–15), and Community Process (16–20). Score each applicable criterion met or not met; two narrow classes drop from the count (§5.2). A few criteria are deliberately fuzzy — “broad economic-node support” is argued from the evidence, not computed from a formula (§5.3).

Read the share of criteria met as a band: **Green** (100%) ready for activation signaling; **Yellow** (75–99%) gaps to close first; **Orange** (50–74%) serious deficiencies; **Red** (below 50%) not ready, with the independent economic-node response on the table (§5.2).

§5.4 works the scorecard through four proposals: **Taproot 17/17 — Green**, **BIP-110 3/20 — Red**, **SegWit2x $\approx$ 5/17 — Red** (the block-size hard fork miners themselves pushed), and a hypothetical rushed pro-data *soft fork* that also scores **Red** despite serving the author’s own fee interest — the cleanest proof the framework measures process, not preference.

Chapter 6

Objections and Responses

6.1 “Bitcoin has no governance.”

Bitcoin has no centralized governance. It has governance. Every consensus change that has ever been adopted required coordination among developers, miners, node operators, and economic actors. That coordination—however informal—is governance. This framework does not propose centralized governance. It proposes minimum standards for evaluating proposals within Bitcoin’s existing decentralized governance structure. The descriptive thesis of *BCAP*—that Bitcoin consensus emerges from the iterated, multi-party interactions of stakeholders with shifting powers and incentives—is correct and is not contested by this framework. The framework’s claim is narrower and consistent with it: minimum standards, public and applied consistently, make the iterated process safer and more efficient by giving every stakeholder a shared vocabulary for turning down proposals that are not ready.

6.2 “Anyone can run whatever software they want.”

True. And this framework does not propose restricting that right. Node operators are free to run any software they choose. This framework proposes that the community develop shared standards for evaluating proposals, so that node operators can make informed decisions. A node operator who runs an activation client that fails every criterion in this framework is exercising their right. They are also assuming quantifiable risks that they may not fully understand. A framework for understanding those risks serves the same function as securities disclosure: it does not restrict choice, it informs it.

6.3 “This framework would prevent necessary changes.”

This framework would slow down reckless changes. It would not prevent necessary ones. SegWit and Taproot both would have passed this framework with high scores. The changes this framework would impede are the ones that should be impeded: poorly reviewed, inadequately tested, rashly activated proposals that put the network at risk.

A sharper version of this objection turns the framework on Bitcoin’s own history: several changes now regarded as beneficial would sit awkwardly against these floors. P2SH was deployed in 2012 around the 55% threshold the framework calls reckless. CLTV and CSV

were low-controversy soft forks that activated on timelines far shorter than the framework's longer review windows. And SegWit's escape from miner deadlock was driven by the threat of BIP-148, the user-activated soft fork the framework treats as a measure of last resort. If the standards would have impeded those, are they not too strict?

The objection confuses the floors with the cases; take them in turn. P2SH is in fact the framework's witness, not the objection's: as §2.1 recounts, its 55% poll fell short and the chain forked before a flag day rescued the change—all of it on a network whose market capitalization was then orders of magnitude smaller than today's. The framework's threshold guards against the risk of a chain split (§3.4), and the cost of a split scales with the value exposed to it: a fork the early network could absorb is not survivable on a trillion-dollar settlement network. The lesson of P2SH is not that 55% is safe—it produced the very split the model predicts—but that the tolerable threshold rises with the stakes, which have risen by orders of magnitude since 2012. CLTV and CSV are a separate point: they reused previously reserved no-op opcodes to add new validation behavior, without restricting any transaction type then in use, which places them in the framework's lower-risk tier (Category 2), and both were technically uncontested. The framework's long review windows are reserved for high-risk and contested changes; a risk-tiered standard holds low-risk, uncontested improvements like these to its most modest floor, not its longest—so they are not the case the floors are built to catch. BIP-148, finally, is not a counterexample to the UASF standard but an instance of it met: SegWit had roughly twenty months of review and broad community support blocked only by miner signaling—the precise conditions under which the framework holds a user-activated soft fork legitimate (§3.4). The framework would have passed BIP-148 and failed BIP-110 for the same reason it distinguishes them today: one carried consensus, the other did not. A standard that endorses the changes the network was right to make and flags the ones it was right to refuse is not too strict. It is calibrated.

Bitcoin's value proposition is stability, predictability, and resistance to arbitrary change. A framework that makes consensus changes harder to execute is aligned with that value proposition, not contrary to it.

6.4 “Who decides whether the standards are met?”

Everyone. And no one. This framework is a tool, not an authority. Any member of the community can evaluate a proposal against these criteria and publish the results. There is no certification body, no approval committee, and no veto power. The framework's authority derives from its usefulness. If the community finds it useful, it will be adopted. If not, it will be ignored. That is how governance works in a decentralized system.

6.5 “The legal analysis is speculative.”

All legal analysis of novel situations is, to some degree, speculative. No court has ruled on the liability of a Bitcoin developer for a chain split caused by a reckless activation. But the absence of precedent does not mean the absence of risk. The legal principles applied in Section 4 — negligence, tortious interference, fiduciary duty — are well established. Their application to Bitcoin governance is novel but not unprecedented. Courts routinely apply existing legal frameworks to new technologies. The question is not whether these principles apply, but how. This paper offers an analysis, not a prediction. And the analysis does not wait on a court to matter: a dated, public score measuring a proposal against a known standard creates notice and foreseeability — the elements these theories turn on — before any harm occurs, which is what makes the risk concrete rather than abstract. The exposure is real the moment the facts are; a precedent would only name what was already there.

A related objection — that a privately written rubric cannot supply a standard of care — proves less than it seems. Courts routinely admit voluntary standards and industry custom as evidence of what reasonable care permits and what precautions were feasible, with weight that grows as adoption grows; and since *The T.J. Hooper*, 60 F.2d 737 (2d Cir. 1932), even universal custom has not capped the care reasonableness demands. The framework does not need adoption to describe the risks accurately. Adoption only raises the evidentiary price of ignoring it.

6.6 “If the standards are not enforceable, what does the framework add?”

This objection has two forms. The strong form: a benign proposal would already meet the framework’s criteria, and an ill-considered proposal will be resisted by the same stakeholders for the same reasons whether the framework exists or not — so the framework adds nothing operational. The weak form: the framework may legitimize rule-lawyering — proposals that satisfy the criteria in form but not in spirit.

The strong form misreads where the framework’s value sits. The value is not in creating coordinated resistance to ill-considered proposals — that coordination exists, however imperfectly, in the form of Bitcoin’s distributed governance. BIP-110’s failure to attract more than a fraction of one percent of voluntary block-signaling demonstrates the existing coordination’s capacity. The framework’s value is in lowering the coordination cost of principled refusal — the cost, that is, of articulating and defending a refusal each operator reaches on its own, not the cost of agreeing with other operators to refuse (§5.2). Without the framework, every individual stakeholder objection requires a custom argument: an exchange operator who refuses to recognize a chain must justify the refusal from first principles, against the rhetorical pressure of proponents who can frame the refusal as arbitrary. With the framework, “this proposal fails criteria 1, 8, 11, and 14” is a justification an objector can state cheaply

and a proponent can answer only by contesting the scoring on its merits—which moves the dispute onto the framework’s ground, where the criteria are public and the scoring is challengeable. Distributed enforcement works, but it works at a cost. The framework lowers that cost without claiming authority it does not have.

This matters most at the margins, where stakeholders are uncertain whether their objection rises to the level of public dissent. A single skeptical infrastructure operator, holding the framework’s criteria, can invoke it without first organizing a coalition. The framework turns what would otherwise be a coordination problem—every stakeholder waiting to see whether others will refuse—into a documentation problem: anyone can point to the criteria and say here is the standard; here is the gap. The credibility of the resulting refusal does not depend on the refuser’s individual standing but on the criteria’s defensibility.

The weak form—that the framework legitimizes rule-lawyering—assumes that a proposal which games the letter of the criteria thereby earns approval. It does not. Binary criteria and classification bands leave judgment intact; partial gaming is itself progress (a rule-lawyered proposal is at least better documented than an ad-hoc one); the framework can be revised adversarially as bad-faith actors expose loopholes; and §5.2’s Red classification remains available for proposals that satisfy the letter but fail the spirit—which is itself observable evidence in the scoring record.

A more honest statement of the weak form is: any standard can be gamed. That is true of every standard humans have ever written. It is not a reason to write none; it is a reason to write standards that produce useful documentation of the gaming when it happens. This framework does that. BIP-110, evaluated against the framework, scores 3/20—a documented Red (§5.4). That documentation, available to any stakeholder, makes a future rule-lawyered variant easier to evaluate against than the original, not harder. The framework’s effect is therefore cumulative: each evaluation makes the next one more efficient, because the prior scoring is in the record.

The harder version of the gaming worry is the active one. A determined, well-funded proponent could manufacture the criteria that are cheap to manufacture—commission three nominally independent reviews, run a testnet deployment as theater, publish a contingency-plan document, and set its activation threshold at the 90% floor—and so assemble a Yellow score around a proposal that has no genuine consensus. The framework’s defense is that the criteria a proponent can manufacture are not the criteria that decide the question. Demonstrated broad economic-node support (criterion 12), no sustained and unaddressed opposition, and adequate community discussion (criterion 17) require the assent of parties the proponent does not control and cannot be faked by spending money. A high score on the manufacturable criteria combined with failure on the assent criteria is not a near-miss; it is itself the diagnostic—a proposal that has assembled the form of readiness without its substance, displayed as such in the scoring record. The fuzzy, judgment-preserving criteria of §5.3 are not the framework’s weakness. They are what a standard that resists gaming requires.

6.7 “This framework is itself a centralizing instrument.”

This is the strongest form of the authority objection, and it deserves a direct answer rather than the dodge that “the framework is just a tool” can become. The objection runs: a privately written twenty-point rubric that tells economic nodes to decline to recognize a chain as Bitcoin (§5.2) is itself an attempt to coordinate a specific governance outcome—the very kind of pressure on a few entities that the framework elsewhere warns against.

The answer is to distinguish what the coordination is organized around. Coordinating around published, challengeable criteria is not equivalent to coordinating around private preference or insider pressure. The criteria are visible; a stakeholder who invokes them exposes the basis of the refusal to scrutiny, and anyone who disagrees can contest the scoring on the record. The framework binds no one, certifies nothing, and confers no veto; its “decline to recognize” step describes an option economic nodes already have, not an instruction they must follow. A centralizing instrument concentrates the power to decide; this framework distributes the power to argue. The two are not the same, and the difference is the whole point. The same design choice keeps the response on the right side of the antitrust line (§5.2): published criteria applied unilaterally are information. Only agreement would make them something else.

6.8 “This framework is a chilling instrument that will deter good-faith development.”

This is the strongest objection a critic can level at the paper itself, and it is sharpest precisely because the author is a practicing litigator. The objection runs: the paper pairs a worked tort map against developers (§4.1) with a scorecard whose Red finding Section 4 frames as “the exposure” (§4.0). That pairing, the objection says, is not a neutral description of risk. It is an instrument of intimidation. Its predictable effect is to chill the open-source contribution Bitcoin depends on, because a contributor who reads it will reasonably fear that any change carrying their name could be turned into a damages theory. Worse, an adversary could weaponize the pairing—publish a self-serving low score against a disfavored developer’s proposal and wave §4.1 as the threatened consequence—converting a tool of evaluation into a tool of harassment.

The objection has real force, and the answer is not to deny the teeth but to locate them precisely. First, the liability theories Section 4 maps are not the paper’s invention. Negligent misrepresentation, tortious interference, and the duty questions left open by *Tulip Trading* exist in the law whether or not anyone writes them down; Section 4 describes the terrain, it does not lay the mines. A contributor faces exactly the exposure the law already assigns—the paper changes the contributor’s information, not the contributor’s risk. Second, and decisively, Section-3 compliance is a strong defense. A change that documents its review, meets its threshold, and clears the floors scores Green and carries no exposure the framework

can manufacture; the net effect on a good-faith contributor is therefore protective, giving them a documented standard of care to point to rather than a trap to fall into. Third, the deterrence the paper theorizes (§4.0) runs against one thing only: shipping a change past a documented Red finding — against recklessness, never against contribution, and never against non-adoption. Good-faith coding that turns out mistaken is not its target; rushing a flagged change to forced activation is. The framework does not deny a chilling effect, then; it embraces a narrow one — deterring the activation of a change already documented as reckless is the intended function, not an unfortunate side effect, and a deterrent that bites only recklessness is the kind a good-faith contributor should want to exist. Fourth, the framework deliberately trains its analysis, and the deterrence it theorizes, on identifiable campaign actors — pools, deploying platforms, conflicted promoters who push past the red flags — not on pseudonymous coders contributing in good faith. The adversary who tries to weaponize a manufactured score inherits the framework’s own discipline: the score is challengeable on the record (§5.3, §6.7), and a score unsupported by the underlying facts documents the harasser, not the target. That challenge has a venue: the dissenting-evaluation archive the Conclusion commits to, where a contested score is met and answered on the record rather than left to assertion. One asymmetry remains: a manufactured Red can be published the instant a proposal appears, before any rebuttal is archived. But a score carries only the weight of the evidentiary basis §5.3 requires it to publish, and a Red asserting failure of the assent criteria (12 and 17) must show the absent support or unaddressed opposition it claims — so a basis-less manufactured Red is the form of a score without its substance (§6.6), self-diagnosing on its face by the very standard it invokes, even before the archive answers it. That answer is epistemic; the framework adds a procedural one for the crudest case. A score published without the criterion-by-criterion evidentiary basis §5.3 requires is not one this framework recognizes as a scorecard — it is an assertion wearing a number — so the target of a bare-number manufactured Red can point to that absence rather than refute it on the facts before the archive answers. A harasser who instead fabricates a plausible basis defeats that shortcut, but not the check beneath it: a basis is only as good as the record it cites, and a fabricated one is refuted on that same public record, the way the archive answers any contested score. The procedural point disposes of the bare-number Red at a glance; the record answers the dressed-up one on the merits.

Chapter 6 in brief

Eight common objections, and the framework's response to each:

- *"Bitcoin has no governance."* It has decentralized governance. The framework gives that process a shared vocabulary, not a central authority (§6.1).
- *"Anyone can run whatever software they want."* True. The framework does not restrict that. It informs the choice (§6.2).
- *"This would prevent necessary changes."* SegWit and Taproot would pass with high scores. The framework impedes reckless changes, not necessary ones (§6.3).
- *"Who decides whether the standards are met?"* Everyone, and no one. The framework is a tool, not an authority (§6.4).
- *"The legal analysis is speculative."* All novel-application legal analysis is. The principles applied are established; the application is new (§6.5).
- *"If the standards aren't enforceable, what does the framework add?"* It lowers the coordination cost of principled refusal — *"this proposal fails criteria 1, 8, 11, and 14"* is a complete justification, cheap to invoke and costly to rebut (§6.6).
- *"This framework is itself a centralizing instrument."* Coordinating around published, challengeable criteria is not the same as concentrating the power to decide — the framework distributes the power to argue, binds no one, and certifies nothing (§6.7).
- *"This framework is a chilling instrument that will deter good-faith development."* The liability theories pre-exist the paper; Section-3 compliance is a strong defense; the deterrence runs only against shipping past a documented Red finding, never against good-faith contribution or non-adoption, and it points at identifiable campaign actors, not pseudonymous coders (§6.8).

Chapter 7

Conclusion

Bitcoin’s consensus rules govern the creation and transfer of value for millions of people and the storage of wealth measured in trillions of dollars. Changes to these rules should be evaluated with a rigor matching their stakes.

The current system—in which proposals are evaluated ad hoc, activation mechanisms are invented on the fly, review periods range from weeks to years with no standard, and evaluation happens in mailing-list threads, social-media arguments, and GitHub comments with no shared benchmark—is inadequate. It has produced near-catastrophic chain splits, wasted years of developer time on governance disputes, and created opportunities to push poorly reviewed changes to activation. BIP-110 attracted negligible voluntary signaling—a peak share under one percent through July 2026—and because it deploys as a LOT=true mechanism, the failed persuasion campaign ends not in withdrawal but in a scheduled confrontation: mandatory signaling at blocks 961,632–963,647 (August 2026), forced activation at block 965,664 (early September 2026). That it never commanded genuine consensus is a fact of community vigilance, not of structural protection. The next proposal of its kind will arrive on the same terms—no required review, no minimum code quality, no agreed-upon threshold—unless this gap is filled. That is not one lawyer’s forecast. Lopp, closing his 2026 retrospective on the block-size war: “as long as Bitcoin is still growing, we should expect more battles over how the protocol should evolve and what trade-offs should be accepted in order to improve the system in different ways.”

This framework does not solve the fundamental challenge of decentralized governance. No framework can. What it provides is a common vocabulary, a shared set of criteria, and a concrete checklist to evaluate proposals against. It shifts the burden of proof onto proponents of change—where it belongs—and provides the community with a structured way to say: this proposal is not ready.

A standard that prescribes process owes its own. This framework is maintained in a public repository; proposed changes proceed by issue and pull request; substantive revisions are batched into numbered editions with a published changelog; and dissenting evaluations—alternative scorings of the same proposal, with their evidentiary bases—are archived alongside the editions, so the record builds up rather than scatters. The revision process is deliberately the same kind the framework asks of consensus proposals: documented, public, and challengeable.

The framework is licensed under Creative Commons Attribution 4.0 International. It may be freely shared, adapted, and built upon by anyone, for any purpose, with attribution. It is available on GitHub for community review and amendment. If it is useful, it will be used. If it can be improved, it should be improved. That is how Bitcoin works. That is how Bitcoin's governance should work too.

Asaf Fulks

Appendix A — Documentary Record for the BIP-110 Case Study

The framework’s case study should meet the framework’s own documentary standard. This appendix indexes the factual assertions of §1.3 to their primary sources. Archived captures of each source—Internet Archive, with capture dates—are maintained in the companion repository at github.com/asafufks/consensus-change-standards, so that the record survives link rot and post-hoc edits. Nothing in this appendix is offered for the truth of the sources’ own claims; it is offered to show what was published, where, and when.

A.1 Initial proposal. Bitcoin development mailing list post introducing the Reduced Data Temporary Softfork, 26 October 2025. Thread: groups.google.com/g/bitcoinddev/c/n0Zim6FbuF8. [Archived: web.archive.org/web/20251028174929/https://groups.google.com/g/bitcoinddev/c/n0Zim6FbuF8, captured 28 October 2025.]

A.2 Number assignment. BIP number 110 assigned 3 December 2025 (bitcoin/bips pull request #2017; editor announcement). [Archived (PR as filed, pre-assignment): web.archive.org/web/20251105193642/https://github.com/bitcoin/bips/pull/2017, captured 5 November 2025; the assigned number appears in the A.3 capture.]

A.3 Specification. bip-0110.mediawiki: 55% signaling threshold; mandatory signaling during blocks 961,632–963,647; LOCKED_IN at height 963,648; ACTIVE at max_activation_height 965,664; expiry approximately one year after activation; 34-byte cap on non-OP_RETURN scriptPubKeys; 83-byte cap on OP_RETURN outputs; tapscript and witness restrictions. github.com/bitcoin/bips/blob/master/bip-0110.mediawiki. [Archived: web.archive.org/web/20260216063723/https://github.com/bitcoin/bips/blob/master/bip-0110.mediawiki, captured 16 February 2026; a later capture, 5 June 2026.]

A.4 Activation client. UASF BIP-110 v0.1 Release Candidate 1, released 10 December 2025 as a fork of Bitcoin Knots (tag v29.2.knots20251110+bip110-v0.1rc1). github.com/dathonohm/bitcoin/releases. [Archived: web.archive.org/web/20251211004020/https://github.com/dathonohm/bitcoin/releases, captured 11 December 2025—one day after release.]

A.5 Defect reports. Developer reports that the client’s test suite—including the client author’s own activation tests—was failing, with warnings that operators running the code could fork themselves off the network (mailing-list and review threads; the specific posts relied on in §1.3, item B). [Archived (full thread including the subsequent defect discussion): web.archive.org/web/20260212202430/https://groups.google.com/g/bitcoinddev/c/n0Zim6FbuF8, captured 12 February 2026.]

A.6 Platform distribution. Node-management platforms listing the activation client as a selectable version alongside stable Knots releases, with no risk disclosure or visual differentiation at the point of selection (§1.3, item D). Umbrel: “BIP110” added as a version option in the Bitcoin Knots app’s settings (getumbrel/umbrel-apps pull request #4237, opened 11 December 2025, merged 22 January 2026; updated by #4720, merged 19 February 2026; neither adds warning or risk language at the point of selection). myNode: BIP-110 added to the Custom Bitcoin Versions list (mynodebtc/mynode pull requests #982, opened 5 January 2026, through #992, merged 17 February 2026). The proponent’s installation guide documents both selection flows. [Archived: installation guide, web.archive.org/web/20260311071550/https://bip110.org/howto/, captured 11 March 2026; the 15 January 2026 capture (web.archive.org/web/20260115224116/https://bip110.org/howto/) shows both listings as pending pull requests; Umbrel PR pages captured 13 January and 9 June 2026; myNode PR pages captured 8 January and 9 June 2026.]

A.7 Default-stream bundling. Bitcoin Knots v29.3.knots20260508 release notes, 9 May 2026: BIP-110 (RDTs) rules included in the stable release; runtime confirmation via `consensusrules=rdts` or GUI dialog; characterization of the change as fixing critical vulnerabilities; assertion that the upgrade already has broad community support; pointer to the one-day-older non-RDTs variant v29.3.knots20260507. github.com/bitcoinknots/bitcoin/blob/v29.3.knots20260508/doc/release-notes.md. [Archived: web.archive.org/web/20260609220415/https://github.com/bitcoinknots/bitcoin/blob/v29.3.knots20260508/doc/release-notes.md, captured 9 June 2026.]

A.8 Discouraged-variant labeling. bitcoinknots.org download page presenting the non-RDTs variant as not recommended. [Archived: web.archive.org/web/20260609220525/https://bitcoinknots.org/, captured 9 June 2026; the prior capture (1 May 2026) predates the v29.3.knots20260508 release.]

A.9 Signaling data. Peak observed share of blocks voluntarily signaling readiness (under one percent through July 2026); tracker: bip110monitor.com. [Archived: web.archive.org/web/20260609220708/https://bip110monitor.com/, captured 9 June 2026; a later capture, web.archive.org/web/20260712175723/https://bip110monitor.com/, 12 July 2026, is contemporaneous with this revision.]

A.10 Advocacy claims. bip110.org: urgency and support claims made in the activation campaign (§1.3; §4.2). [Archived: web.archive.org/web/20260106225951/https://bip110.org/, captured 6 January 2026; later captures 24 January and 10 March 2026.]

Appendix B — Adoption Kit: Model Policy Language

The framework becomes useful the day an operator can adopt it without drafting from scratch. The clauses below are starting points, written to be applied unilaterally — each operator to its own independent judgment — consistent with §5.2’s antitrust note. They are illustrative drafting for adaptation by qualified counsel, not legal advice, and they create no obligation to act in concert with any other party. An operator adopting any clause below cites the readiness criteria (§5.1) as one risk-management input to its own decision; it takes no position on, and assumes no role in, the liability analysis of Section 4, which addresses a proponent’s exposure for shipping past a documented finding, not the operator’s recognition decision. A standalone scorecard (`scorecard.md` and a fillable PDF) is maintained in the companion repository at github.com/asaffulks/consensus-change-standards.

B.1 Recognition policy (exchange / custodian). “[Operator] evaluates proposed changes to the Bitcoin consensus rules against the Consensus Change Readiness Checklist (Fulks, *Consensus Change Standards* (5th ed. 2026), §5.1) as one input to its independent listing, custody, deposit-crediting, and contract-denomination decisions. A proposal that has not been publicly scored, or that scores below the Yellow band, weighs heavily against recognition as ‘Bitcoin’ for those purposes in [Operator]’s own independent assessment, which [Operator] makes for itself and may revisit as the proposal’s record develops. This policy is adopted and applied unilaterally; the published score is one evidentiary input to [Operator]’s judgment, not a substitute for it, and this policy neither requires nor contemplates agreement or coordinated action with any other market participant.”

B.2 Signaling policy (mining pool). “[Pool] will not signal readiness for any consensus change that has not completed the minimum review period for its risk category (§3.2) and published a chain-split risk assessment (§3.5). [Pool] discloses each signaling decision, and the evaluation supporting it, to connected miners at least thirty days before signaling begins, and provides a mechanism for miners to redirect hashrate without penalty during that period.”

B.3 Distribution policy (node-software platform). “[Platform] lists software that alters Bitcoin consensus rules separately from maintenance releases, labels it as consensus-altering at the point of selection, presents a plain-language risk disclosure before installation, and does not include consensus-altering defaults in any update path presented as routine.”

B.4 Citation in adopted policies. Operators adopting any clause above should cite the edition and version evaluated against, so that scorings remain comparable as the framework is revised.

Glossary of Technical Terms

Activation threshold.

The percentage of mining hashrate that must signal readiness for a soft-fork change before the change locks in. (Signaling at lock-in is distinct from enforcement at activation; see the 55% discussion in §3.4.) Historically 95% (SegWit via BIP-9) and 90% (Taproot via Speedy Trial); BIP-110 proposed 55%. See §3.4.

BCAP.

Shortened form for Ren Crypto Fish, Steve Lee & Lyn Alden, *Analyzing Bitcoin Consensus: Risks in Protocol Upgrades* (Nov. 2024). The principal prior analytical work this paper builds on. See §1.5.

BIP (Bitcoin Improvement Proposal).

A formal proposal document for a change to the Bitcoin protocol or surrounding standards. Filing a BIP does not activate the change; activation is a separate ecosystem process this framework evaluates.

Bitcoin Core.

The reference implementation of the Bitcoin protocol, maintained by the Bitcoin Core project. The codebase against which other implementations are typically measured for consensus compatibility.

Bitcoin Knots.

A node implementation derived from Bitcoin Core with additional policy options. The BIP-110 activation client is a fork of Bitcoin Knots; the v29.3.knots20260508 release bundled BIP-110 (RDTs) activation rules into the default release stream. See §1.3.

Chain split.

A divergence in the network where some nodes accept a block under one set of consensus rules and other nodes accept a different block under different rules, producing two parallel chains.

Consensus rules.

The network-wide rules that determine whether a block is valid. Changes to consensus rules require either a soft fork or a hard fork.

Economic Nodes.

Full nodes operated by entities with material Bitcoin throughput — exchanges, custodians, payment processors, large merchants, ETF operators. Their enforcement of consensus rules is what gives miner signaling its weight.

Hard fork.

A consensus-rule change that loosens or alters validity rules such that the new rules accept blocks the old rules would reject. Pre-upgrade nodes cannot follow the new chain. See §3.7.

Hashrate.

The aggregate computational power miners are dedicating to Bitcoin’s proof-of-work. Used as the unit of measure for activation signaling.

Inscriptions.

Image, text, document, or binary files embedded in the witness data of Bitcoin transactions using Casey Rodarmor’s Ordinals protocol. The protocol encodes data within an unexecuted `OP_FALSE OP_IF . . . OP_ENDIF` envelope inside Taproot’s script-path spend; SegWit’s witness-data fee discount makes this storage substantially cheaper, per byte, than alternatives such as `OP_RETURN`. The spread of inscriptions since the protocol’s January 2023 release is the motivating context for the wave of restrictive consensus-change proposals discussed in this paper. See §1.2.

MASF (Miner-Activated Soft Fork).

A soft fork that activates when miner signaling crosses the activation threshold. The default activation pathway in Bitcoin.

OP_RETURN.

A Bitcoin script opcode that marks a transaction output as provably unspendable, used to embed arbitrary data in a transaction. BIP-110 sought to restrict the size of `OP_RETURN` payloads alongside other forms of data embedding. See §1.3.

OP_SUCCESS.

Opcodes reserved during the Taproot soft fork as placeholders for future additions. Substituting an `OP_SUCCESS` slot with a new opcode enables adding new spending paths via soft fork. See §3.5.

Reference implementation.

The canonical software implementing a proposed consensus change, against which other client implementations are compared for correctness.

Relay policy (standardness).

Node-level rules governing which unconfirmed transactions a node accepts into its mempool and relays to peers, as distinct from the consensus rules that govern which blocks are valid. Relay policy is a local default each operator can change without affecting consensus: a transaction that violates a node’s relay policy can still be mined, and once in a block must be accepted by that node. The distinction is load-bearing in §1.2 — Bitcoin Core v30’s change to its `OP_RETURN` data-carriage default altered relay policy and bound no one, whereas a consensus change alters the definition of a valid block for every participant. See §1.2, §3.2.

Reorganization (reorg).

When a node switches its best chain from one valid tip to a competing chain of greater accumulated work, undoing the transactions in the abandoned blocks. The mechanism by which a sub-overwhelming-enforcement soft fork can be unwound. See §3.4.

Replay protection.

A technical mechanism (typically a transaction-format change) that stops a transaction valid on one chain after a split from being broadcastable on the other. Necessary in hard forks where a meaningful minority continues to run the prior rules. See §3.5.C, §3.7.

Signaling.

Miners indicating readiness for a proposed soft fork by setting a designated bit in block headers during a defined signaling window.

Soft fork.

A consensus-rule change that tightens validity rules such that every block valid under the new rules remains valid under the old (the set of blocks valid under the new rules is a strict subset of those valid under the old). Pre-upgrade nodes continue to follow the upgraded chain so long as a majority of hashrate enforces the new rules (§3.4 models the failure case).

Speedy Trial.

An activation deployment pattern with a fixed signaling window, a high activation threshold (commonly 90%), and a clean timeout if the signaling threshold is not met. Speedy Trial was implemented as a bounded BIP-9 deployment rather than under BIP-8, but its clean timeout on failure — a plain expiry with no force-activation — behaves like BIP-8's LOT=false fail-safe; the alternative, LOT=true, forces activation at the deadline regardless of signaling. See §3.4 on the LOT debate. Used for Taproot in 2021.

Sunset clause.

A provision that makes a consensus rule change automatically expire at a defined block height or median time past, returning the network to pre-activation rules without further intervention. A valid sunset is self-executing (no further software action required) and tested on testnet to confirm deactivation works correctly. See §3.6.

UASF (User-Activated Soft Fork).

A soft fork activated by economic nodes enforcing new rules without depending on a miner-signaling threshold being met. BIP-148 is the canonical example. A LOT=true deployment is a hybrid form — it sets a miner-signaling threshold but forces activation at timeout regardless, so it operates as a UASF when miners do not comply; BIP-110 is of this kind. UASF works only when prior community support is overwhelming; UASF without that support yields the BIP-110 pattern — negligible voluntary signaling and a forced-activation confrontation. See §3.4.

UTXO (Unspent Transaction Output).

The fundamental unit of Bitcoin ownership: a transaction output that has been received but

not yet spent. The complete set of UTXOs at a given block height — the “UTXO set” — defines all spendable Bitcoin at that moment.

References

Bitcoin Improvement Proposals

- BIP-1: *BIP Purpose and Guidelines*. Amir Taaki, 2011. github.com/bitcoin/bips/blob/master/bip-0001.mediawiki
- BIP-2: *BIP Process, Revised*. Luke Dashjr, 2016 (replaced by BIP-3, 2025). github.com/bitcoin/bips/blob/master/bip-0002.mediawiki
- BIP-3: *Updated BIP Process*. Murch (Mark Erhardt), 2025 (Status: Deployed; replaces BIP-2). github.com/bitcoin/bips/blob/master/bip-0003.md
- BIP-8: *Version Bits with Lock-in by Height*. Shaolin Fry, Luke Dashjr, 2017. github.com/bitcoin/bips/blob/master/bip-0008.mediawiki
- BIP-9: *Version Bits with Timeout and Delay*. Pieter Wuille, Peter Todd, Greg Maxwell, Rusty Russell, 2015. github.com/bitcoin/bips/blob/master/bip-0009.mediawiki
- BIP-12: *OP_EVAL*. Gavin Andresen, 2011. github.com/bitcoin/bips/blob/master/bip-0012.mediawiki
- BIP-16: *Pay to Script Hash*. Gavin Andresen, 2012. github.com/bitcoin/bips/blob/master/bip-0016.mediawiki
- BIP-17: *OP_CHECKHASHVERIFY (CHV)*. Luke Dashjr, 2012. github.com/bitcoin/bips/blob/master/bip-0017.mediawiki
- BIP-50: *March 2013 Chain Fork Post-Mortem*. Gavin Andresen, 2013. github.com/bitcoin/bips/blob/master/bip-0050.mediawiki
- BIP-65: *OP_CHECKLOCKTIMEVERIFY*. Peter Todd, 2014; activated 2015. github.com/bitcoin/bips/blob/master/bip-0065.mediawiki
- BIP-68: *Relative lock-time using consensus-enforced sequence numbers*. Mark Friedenbach, BtcDrak, Nicolas Dorian, kinoshitajona, 2015. github.com/bitcoin/bips/blob/master/bip-0068.mediawiki
- BIP-91: *Reduced threshold Segwit MASF*. James Hilliard, 2017. github.com/bitcoin/bips/blob/master/bip-0091.mediawiki
- BIP-101: *Increase Maximum Block Size*. Gavin Andresen, 2015. github.com/bitcoin/bips/blob/master/bip-0101.mediawiki
- BIP-102: *Block Size Increase to 2MB*. Jeff Garzik, 2015. github.com/bitcoin/bips/blob/master/bip-0102.mediawiki

- BIP-110: *Reduced Data Temporary Softfork* (originally proposed as BIP-444; number assigned 3 December 2025). Dathon Ohm, 2025. github.com/bitcoin/bips/blob/master/bip-0110.mediawiki
- BIP-112: *CHECKSEQUENCEVERIFY*. BtcDrak, Mark Friedenbach, Eric Lombrozo, 2015. github.com/bitcoin/bips/blob/master/bip-0112.mediawiki
- BIP-113: *Median time-past as endpoint for lock-time calculations*. Thomas Kerin, Mark Friedenbach, 2015. github.com/bitcoin/bips/blob/master/bip-0113.mediawiki
- BIP-141: *Segregated Witness (Consensus Layer)*. Eric Lombrozo, Johnson Lau, Pieter Wuille, 2015. github.com/bitcoin/bips/blob/master/bip-0141.mediawiki
- BIP-148: *Mandatory Activation of Segwit Deployment*. Shaolin Fry, 2017. github.com/bitcoin/bips/blob/master/bip-0148.mediawiki
- BIP-340: *Schnorr Signatures for secp256k1*. Pieter Wuille, Jonas Nick, Tim Ruffing, 2020. github.com/bitcoin/bips/blob/master/bip-0340.mediawiki
- BIP-341: *Taproot: SegWit Version 1 Spending Rules*. Pieter Wuille, Jonas Nick, Anthony Towns, 2020. github.com/bitcoin/bips/blob/master/bip-0341.mediawiki
- BIP-342: *Validation of Taproot Scripts*. Pieter Wuille, Jonas Nick, Anthony Towns, 2020. github.com/bitcoin/bips/blob/master/bip-0342.mediawiki
- BIP-361: *Post Quantum Migration and Legacy Signature Sunset*. Jameson Lopp, Christian Papathanasiou, Ian Smith, Joe Ross, Steve Vaile, Pierre-Luc Dallaire-Demers, 2026. github.com/bitcoin/bips/blob/master/bip-0361.mediawiki

Legal Authorities

- *Aas v. Superior Court*, 24 Cal.4th 627 (2000).
- *Allied Tube & Conduit Corp. v. Indian Head, Inc.*, 486 U.S. 492 (1988).
- *Bell Atlantic Corp. v. Twombly*, 550 U.S. 544 (2007).
- *Biakanja v. Irving*, 49 Cal.2d 647 (1958).
- *Bily v. Arthur Young & Co.*, 3 Cal.4th 370 (1992).
- *Brown v. USA Taekwondo*, 11 Cal.5th 204 (2021).
- *Carma Developers (Cal.), Inc. v. Marathon Development California, Inc.*, 2 Cal.4th 342 (1992).
- *Crypto Open Patent Alliance v Wright* [2024] EWHC 1198 (Ch), High Court of England and Wales (Ch. Div.) (Mellor J, 20 May 2024) (holding that Dr. Craig Wright is not Satoshi Nakamoto).
- *Della Penna v. Toyota Motor Sales, U.S.A., Inc.*, 11 Cal.4th 376 (1995).
- *E.I. du Pont de Nemours & Co. v. FTC*, 729 F.2d 128 (2d Cir. 1984).

- *FTC v. Indiana Federation of Dentists*, 476 U.S. 447 (1986).
- *FTC v. Superior Court Trial Lawyers Ass’n*, 493 U.S. 411 (1990).
- *Guz v. Bechtel National, Inc.*, 24 Cal.4th 317 (2000).
- *Interstate Circuit, Inc. v. United States*, 306 U.S. 208 (1939).
- *J’Aire Corp. v. Gregory*, 24 Cal.3d 799 (1979).
- *Junger v. Daley*, 209 F.3d 481 (6th Cir. 2000).
- *Klor’s, Inc. v. Broadway-Hale Stores, Inc.*, 359 U.S. 207 (1959).
- *Korea Supply Co. v. Lockheed Martin Corp.*, 29 Cal.4th 1134 (2003).
- *Lemmon v. Snap, Inc.*, 995 F.3d 1085 (9th Cir. 2021).
- *Northwest Wholesale Stationers, Inc. v. Pacific Stationery & Printing Co.*, 472 U.S. 284 (1985).
- *Pacific Gas & Electric Co. v. Bear Stearns & Co.*, 50 Cal.3d 1118 (1990).
- *Quelimane Co. v. Stewart Title Guaranty Co.*, 19 Cal.4th 26 (1998).
- *Rattagan v. Uber Technologies, Inc.*, 17 Cal.5th 1 (2024).
- *Robinson Helicopter Co. v. Dana Corp.*, 34 Cal.4th 979 (2004).
- *Rowland v. Christian*, 69 Cal.2d 108 (1968).
- *Sheen v. Wells Fargo Bank, N.A.*, 12 Cal.5th 905 (2022).
- *Southern California Gas Co. v. Superior Court (S. Cal. Gas Leak Cases)*, 7 Cal.5th 391 (2019).
- *The T.J. Hooper*, 60 F.2d 737 (2d Cir. 1932).
- *Third Story Music, Inc. v. Waits*, 41 Cal.App.4th 798 (1995).
- *Tulip Trading Ltd v van der Laan* [2023] EWCA Civ 83, Court of Appeal of England and Wales (3 February 2023) (claim discontinued 2024 without trial). caselaw.nationalarchives.gov.uk/ewca/civ/2023/83
- *United States v. Colgate & Co.*, 250 U.S. 300 (1919).
- *Wolf v. Walt Disney Pictures & Television*, 162 Cal.App.4th 1107 (2008).
- 15 U.S.C. § 1 (Sherman Act § 1).
- 15 U.S.C. § 45 (Federal Trade Commission Act § 5).
- 47 U.S.C. § 230.
- Rev. Rul. 2019-24, 2019-44 I.R.B. 1004 (Oct. 9, 2019) (tax treatment of cryptocurrency hard forks). irs.gov/pub/irs-drop/rr-19-24.pdf
- *Application of the Federal Securities Laws to Certain Types of Crypto Assets and Certain Transactions Involving Crypto Assets*, 91 Fed. Reg. 13,714 (Mar. 23, 2026) (SEC–CFTC joint interpretive release).

- Restatement (Second) of Torts § 449 (Am. Law Inst. 1965).
- Restatement (Second) of Torts § 552 (Am. Law Inst. 1977).
- Restatement (Third) of Torts: Liability for Economic Harm § 5 (Am. Law Inst. 2020).
- Haque, Raina S., Rodrigo Seira, Brent A. Plummer & Nelson M. Rosario. “Blockchain Development and Fiduciary Duty.” 2 Stan. J. Blockchain L. & Pol’y 139 (2019).
- Walch, Angela. “Deconstructing Decentralization: Exploring the Core Claim of Crypto Systems.” In *Cryptoassets: Legal, Regulatory, and Monetary Perspectives*, edited by Chris Brummer, 39–68. Oxford: Oxford University Press, 2019.
- Walch, Angela. “In Code(rs) We Trust: Software Developers as Fiduciaries in Public Blockchains.” In *Regulating Blockchain: Techno-Social and Legal Challenges*, edited by Philipp Hacker, Ioannis Lianos, Georgios Dimitropoulos & Stefan Eich, 58–81. Oxford: Oxford University Press, 2019.

Comparative Common-Law and EU Authorities

Cited by name in §4.6’s comparative note at doctrinal level; specific reporter citations omitted in keeping with the comparative scope. Practitioners requiring precise references should consult the cited courts’ official databases (Bailii for UK, AustLII for Australia, CanLII for Canada, EUR-Lex for EU).

- *Caparo Industries plc v Dickman* (House of Lords, 1990). Leading modern UK duty-of-care case; three-stage test (foreseeability, proximity, fair-just-reasonable).
- *Hedley Byrne & Co Ltd v Heller & Partners Ltd* (House of Lords, 1964). Foundational UK case on negligent misstatement and assumption of responsibility.
- Property (Digital Assets etc) Act 2025 (UK), 2025 c. 29. Statute (in force December 2025) whose single operative section confirms that a thing is not prevented from being an object of personal property rights merely because it is neither a thing in possession nor a thing in action; extends to England, Wales, and Northern Ireland.
- *Sullivan v Moody* (High Court of Australia, 2001). Australian duty-of-care case departing from the *Caparo* three-stage test.
- *Anns v Merton London Borough Council* (House of Lords, 1978). Predecessor two-stage UK duty-of-care test; overruled in the UK but retained in modified form in Canada.
- *Cooper v Hobart* (Supreme Court of Canada, 2001). Leading Canadian duty-of-care case refining the *Anns* two-stage approach.
- Directive (EU) 2024/2853 on liability for defective products (European Union, adopted 2024), replacing Directive 85/374/EEC. Extends strict product liability to software; excludes free and open-source software developed or supplied outside a commercial activity

(Art. 2(2); Recital 14). Entered into force 8 December 2024; Member State transposition deadline 9 December 2026.

- Swiss Code of Obligations (*Obligationenrecht*) Article 41. General civil-law tort liability clause.

Industry and Academic Analysis

- Bier, Jonathan. *The Blocksize War: The Battle Over Who Controls Bitcoin's Protocol Rules*. Independently published, March 2021.
- Buterin, Vitalik. "Notes on Blockchain Governance." December 17, 2017. vitalik.eth.limo/general/2017/12/17/voting.html
- Crypto Fish, Ren, Steve Lee, and Lyn Alden. *Analyzing Bitcoin Consensus: Risks in Protocol Upgrades*. November 2024. github.com/bitcoin-cap/bcap; also available at bitcoinnews.ch/wp-content/uploads/2024/11/bcap_v1.0.pdf.
- De Filippi, Primavera, and Aaron Wright. *Blockchain and the Law: The Rule of Code*. Cambridge, MA: Harvard University Press, 2018.
- Eyal, Ittay, and Emin Gün Sirer. "Majority Is Not Enough: Bitcoin Mining Is Vulnerable." In *Financial Cryptography and Data Security 2014*, 436–454. Springer, 2014. Also arxiv.org/abs/1311.0243.
- Grunspan, Cyril, and Ricardo Pérez-Marco. "Double Spend Races." *International Journal of Theoretical and Applied Finance* 21, no. 8 (2018). arxiv.org/abs/1702.02867
- Lopp, Jameson. *Reflections Upon Bitcoin's Block Size Debate*. July 1, 2026. blog.lopp.net/reflections-upon-bitcoins-block-size-debate/
- Lopp, Jameson. *When Do Bitcoin Node Operators Upgrade?* blog.lopp.net/when-do-bitcoin-node-operators-upgrade/
- Nakamoto, Satoshi. *Bitcoin: A Peer-to-Peer Electronic Cash System*. 2008. bitcoin.org/bitcoin.pdf
- Narayanan, Arvind, Joseph Bonneau, Edward Felten, Andrew Miller, and Steven Goldfeder. *Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction*. Princeton: Princeton University Press, 2016.
- Rosenfeld, Meni. *Analysis of Hashrate-Based Double Spending*. 2014. arxiv.org/abs/1402.2009
- Voell, Zack. "Ethereum Classic Hit by Third 51% Attack in a Month." *CoinDesk*, August 29, 2020. coindesk.com/markets/2020/08/29/
- Zamfir, Vlad. "Against On-Chain Governance." December 1, 2017. medium.com/@Vlad_Zamfir/against-on-chain-governance-a4ceacd040ca

Open-Protocol Governance

- Bradner, S. *The Internet Standards Process—Revision 3*. RFC 2026, Internet Engineering Task Force, October 1996. rfc-editor.org/rfc/rfc2026
- Housley, R., D. Crocker, and E. Burger. *Reducing the Standards Track to Two Maturity Levels*. RFC 6410, Internet Engineering Task Force, October 2011. rfc-editor.org/rfc/rfc6410
- Resnick, P. *On Consensus and Humming in the IETF*. RFC 7282, Internet Engineering Task Force, June 2014. rfc-editor.org/rfc/rfc7282
- Warsaw, Barry, Jeremy Hylton, David Goodger, and Alyssa Coghlan. *PEP Purpose and Guidelines*. PEP 1, Python Software Foundation, 2000 (rev.). peps.python.org/pep-0001
- *EIP Purpose and Guidelines*. EIP-1, Ethereum, 2015 (Living). eips.ethereum.org/EIPS/eip-1

Software Releases

- Bitcoin Knots v29.3.knots20260508 (RDTs-enabled default release). Bitcoin Knots Project, 9 May 2026. bitcoinknots.org/files/29.x/
- Bitcoin Knots v29.3.knots20260507 (final pre-RDTs release; non-RDTs variant). Bitcoin Knots Project, 8 May 2026. bitcoinknots.org/files/29.x/
- RDTs Activation Client v0.1rc1 (BIP-110 activation client; fork of Bitcoin Knots, tag v29.2.knots20251110+bip110-v0.1rc1). 10 December 2025. github.com/dathonohm/bitcoin/releases

License

- Creative Commons Attribution 4.0 International (CC BY 4.0). creativecommons.org/licenses/by/4.0/