Liquid Staking and the Control-Exposure Wedge*

Marc Dordal i Carreras† Gloria Christina Heesen‡

Kohei Kawaguchi§

April 5, 2026

Abstract

Proof-of-stake deters attacks by keeping validator stake exposed to slashing and depreciation losses. Liquid staking lets operators obtain voting power using pooled stake while reducing their own exposure by selling liquid staking tokens (LSTs) and shifting uncovered slashing losses onto token holders. We study the security implications of this control-exposure wedge and the protocol design problem it creates. Competitive LST pricing can partly deter attack by lowering the resale value of claims when risk rises, but it cannot fully restore deterrence because liquid staking participants do not internalize ETH-wide depreciation losses. A fee-charging protocol prefers the no-attack regime because it maximizes total stake, yet collateral requirements alone do not generally make that outcome unique. Robust security may therefore require additional tools, including permissioned participation, screening or reserve capacity.

Keywords: proof-of-stake, liquid staking, blockchain security, decentralized finance

JEL codes: D47, D86, G23, G32

*We thank the participants at the University of Tokyo Game Theory Workshop 2026 and the HKUST CEP brownbag seminar for helpful comments and suggestions.

†The Hong Kong University of Science and Technology. Email: marcdordal@ust.hk.

‡The Hong Kong University of Science and Technology. Email: gcheesen@connect.ust.hk.

§The Hong Kong University of Science and Technology. Email: kkawaguchi@ust.hk.

1 Introduction

Proof-of-stake (PoS) blockchains secure consensus by requiring validators, the entities that propose and attest blocks, to lock the native token as collateral, or "stake."[1] If they violate the protocol rules, that stake can be slashed.[2] If an attack also undermines confidence in the chain, the value of the exposed stake may fall as well. Under the traditional PoS implementation, a validator must keep its full stake position exposed throughout its participation in the consensus protocol. Control and loss-bearing therefore remain aligned (Buterin, 2014; Saleh, 2021).

Liquid staking changes this arrangement. A liquid staking protocol pools user deposits, delegates them to operators who run validators within the traditional PoS mechanism on behalf of the pool and issues liquid staking tokens (LSTs), which are tradable claims on pooled stake net of the portion kept locked as operator collateral. This gives depositors liquidity over their capital while allowing them to delegate specialized validation tasks and continue earning staking rewards. It also creates a new security problem when the same actor can obtain operational control through the pool and then sell the liquid claim before deciding whether to attack. Operators are required to post collateral with the liquid staking protocol, sometimes called a "bond," equal to a fraction of the pooled stake they operate. If slashing losses exceed that collateral, the remaining losses are borne by current LST holders through a decline in the backing value of each token. Liquid staking can therefore separate operational control from economic exposure. We refer to this separation as the control-exposure wedge.

This paper studies the security implications of that wedge. We ask whether rational LST pricing can deter attack and which losses the price can internalize. We also ask whether a fee-charging liquid staking protocol prefers to eliminate the resulting attack risk and whether collateral requirements alone can implement that outcome.

[1] Attesting is effectively casting a vote in the consensus process. In proof-of-stake, voting power is proportional to the amount of stake a validator controls.

[2] Slashing refers to protocol-imposed penalties for clearly attributable misbehavior, such as equivocation. It reduces the validator's effective stake and may also force exit, depending on the protocol design.

To fix ideas, we use Ethereum and its native token ETH as the running example because liquid staking is especially important on Ethereum, the largest proof-of-stake blockchain to date (Gogol et al., 2024a). The mechanism is not Ethereum-specific, however. The same logic applies more broadly, with the exact implications depending on the protocol's attack threshold, slashing rules and staking architecture.

We begin with a two-date model featuring an exogenous pool of honest operators, that is, operators who do not attack consensus regardless of economic incentives,[3] and one strategic operator. At date 1, operators stake ETH through the liquid-staking protocol, keep part of the position locked as protocol-mandated collateral, and sell claims on the remainder in competitive secondary markets. At date 2, the strategic operator, having obtained voting power through the pool, chooses whether to attack. We refer to this as a discount model because the liquid claim is sold at date 1 against stake that can only be redeemed at date 2. Buyers therefore value it as a claim on future backing, denominated in ETH, rather than as spot ETH. Its price can therefore fall below one-for-one in ETH both because slashing risk reduces expected future backing and because buyers may require a positive return.[4]

This setup makes the attacker's trade-off transparent. By selling the LST before the attack decision, the strategic operator remains exposed only through the collateral left at risk, which remains subject to slashing and depreciation in the attack state. Collateral absorbs slashing losses first, and any residual loss is socialized across current LST holders through lower backing per token at date 2. We assume that an attack yields a positive private payoff to the attacker while depressing the value of ETH more broadly. The idea is that a successful attack undermines confidence in the chain's security and settlement reliability, reducing the value of ETH held both inside and outside the liquid-staking pool.

[3] One can think of honest operators as small agents that do not have the stake capacity to mount an attack on their own or, alternatively, as agents whose attack decisions are motivated by considerations other than profit maximization.

[4] We use the discount model as the baseline because it yields sharper closed-form results and makes the pricing, participation, and deterrence channels more transparent. The online appendix develops a compounding version, closer to the practical implementation of major protocols such as Lido, in which users enter the pool at par and hold rebasing claims on pooled stake. Rewards and losses are then transmitted through changes in the claim balance or redemption value rather than through the operator's discounted resale of the LST claims. The main qualitative conclusions are unchanged.

The baseline model has two key objects. The first is the attacker's incentive condition, which determines whether an attack is privately profitable given the price at which the operator can liquidate the LST and the private benefit from attacking. The second is the LST's backing value at date 2, after slashing losses have been allocated between operator collateral and the pool. Under rational pricing, the date-1 LST price equals expected backing, so prices and attack incentives must be mutually consistent. This gives rise to three possible outcomes: no attack, attack and a mixed region in which the attack probability adjusts to make the attacker indifferent.

Price adjustment partly offsets the control-exposure wedge. When attack risk rises, the LST trades at a discount, which lowers the attacker's resale proceeds. But even fully rational LST pricing cannot restore full deterrence. Because the LST is denominated in ETH, it can reflect losses inside the pool but it cannot internalize the broader decline in ETH's value borne by all native-token holders. The key externality is therefore the combination of liquid exit, socialized slashing and ETH-wide depreciation. Liquid staking weakens deterrence relative to traditional staking because it allows the attacker to offload part of the position before the attack decision, maximizing the voting power obtained per unit of capital committed to the attack.

We then endogenize protocol scale. Honest-operator participation determines how much stake enters the protocol, while LST buyers price the claim relative to available outside returns. Because staking returns on Ethereum decline as total stake increases, operator entry, token pricing and attack incentives are jointly determined. The model therefore admits the same no-attack, attack and mixed outcomes as before, along with the possibility that the liquid-staking protocol shuts down because participation falls to zero.

This endogenous-scale model shows why collateral has no simple effect on protocol scale. Higher collateral protects LST holders by reducing the slashing loss passed on to the pool, which tends to support participation. But it also leaves a larger share of operator value exposed to the post-attack drop in ETH and, when collateral does not fully cover slashing, weakens the price discount that would otherwise help deter attack. As a result, the

same collateral policy can support a large protocol with low attack risk, because stronger participation makes the required attack stake larger, or a small protocol with high attack risk, because weaker participation makes the attack easier to mount.

Finally, we study protocol design. The protocol chooses the collateral requirement and the fee on staking rewards to maximize revenue from operated stake. This objective is narrower than social welfare because the protocol does not internalize the ETH-wide losses an attack imposes on native token holders. Even so, whenever feasible, the protocol prefers the no-attack outcome. The reason is simple: attacks reduce honest participation and thereby shrink the stake base on which the protocol earns fees. If an attack occurs, the protocol also stops earning fee revenue on the attacking stake itself. The protocol therefore does not need to internalize the full ETH-wide damage to dislike attacks. The erosion of its own fee base is sufficient. The challenge is therefore one of implementation rather than preference. Collateral can reduce the set of attack-feasible equilibria, but it need not uniquely select the no-attack equilibrium, even when operator collateral fully absorbs slashing losses in the event of an attack. Robust security may therefore require additional tools, such as permissioned participation, screening, reserve capacity or other institutional safeguards.

This paper contributes to the literature on blockchain security in proof-of-stake systems, where security is an equilibrium outcome shaped by validator incentives, token value and staking participation (Saleh, 2021; John et al., 2021; Jermann, 2024; Cong et al., 2025). LSTs add a further layer because they can separate voting power from loss-bearing.[5] In Ethereum, these claims are issued by protocols built on top of the base layer,[6] but they can still change the mapping between stake, voting power and security of the whole consensus layer.

The closest papers are Tzinas and Zindros (2023), which study liquid staking through a

[5]LSTs expanded rapidly after Ethereum's 2022 transition to proof-of-stake (the "Merge"). Decentralized finance (DeFi) refers to smart-contract-based financial applications on public blockchains. See Heimbach et al. (2023) for evidence on how the Merge changed liquidity and lending activity in DeFi.

[6]The base layer is the underlying blockchain that provides transaction ordering, settlement and consensus. Protocols built on top are smart-contract systems such as liquid staking protocols that pool deposits and issue receipt-like claims.

principal-agent lens and show how fungibility and pooled delegation can socialize slashing losses, and Lehar et al. (2025), which show that secondary markets can both ease illiquidity and coordinate runs through prices. Together, these papers show how liquid staking can generate loss socialization or price-coordinated fragility. We instead focus on proof-of-stake security and model liquid staking as creating a control-exposure wedge. Our approach does not rely on a particular delegation rule, run mechanism, or trading strategy, and shows that liquid staking can weaken deterrence even under rational pricing.

Existing work on staking pools studies the efficiency-security trade-off in delegation and the possibility that malicious actors attract delegated stake (Gersbach et al., 2022). Related empirical and theoretical work links pool competition and governance to concentration dynamics (Jeong, 2020; Irresberger and Yang, 2023; Tang et al., 2024). Our focus is different. When liquid claims on pooled stake are tradable, an attacker can accumulate operational control without bearing the full losses from misbehavior. Price adjustment can partly deter attack, but it cannot internalize ETH-wide depreciation borne by native token holders.

Recent work also maps the design space of liquid staking and restaking[7] and studies how protocol design affects risk, token performance and capital efficiency (Gogol et al., 2024a,b; Carre and Gabriel, 2024). We complement that literature by treating pricing and incentives as a single equilibrium object. In the baseline model, the analysis ties the LST price to the profitability of misbehavior and characterizes the resulting no-attack, attack and mixed regions. The endogenous-scale extension also characterizes shutdown outcomes and studies the protocol's design problem.

The mechanism is also related to finance settings where control is separated from exposure or claims are transferred before losses are realized (Pennacchi, 1988; Bebchuk et al., 2000; Chemla and Hennessy, 2014). In both cases, control can be separated from loss-bearing. In our setting, the security consequence is that even fully rational LST pricing cannot internalize the ETH-wide losses borne by native token holders, nor can it always prevent an increase in attack risk.

[7] Restaking reuses collateral that is already staked on the base chain to provide additional security guarantees to other protocols, usually in exchange for additional fees.

The remainder of the paper is organized as follows. Section 2 provides institutional background on liquid staking and summarizes the main protocol designs. Section 3 studies the discount model with an exogenous operator pool and derives its pricing, equilibrium, and security implications. Section 4 extends the analysis to endogenous operator participation and examines equilibrium properties and the protocol's design problem. Section 5 concludes. The appendix describes the construction of the data used in Section 2 and collects the proofs.

2 Institutional background

Proof-of-stake attacks require controlling a sufficient fraction of actively staked voting power. An entity with a majority can censor transactions or induce chain reorganizations (Ethereum Foundation, 2026a). Protocols with explicit finality rules require a qualified majority of validators to ratify each block.[8] Under Ethereum's consensus rules, this threshold is two-thirds of staked voting power, and an adversary with at least one-third can prevent finality (Ethereum Foundation, 2026c,b).

Liquid staking protocols pool deposits and delegate the resulting voting power to node operators, with operational control separated from the economic claims of depositors. Protocols vary in how operators are admitted, from curated or governance-gated sets to permissionless participation, and this distinction matters for how easily an entity can accumulate operated stake (Lido Finance, 2025b, 2026b).

Capital requirements limit how much voting power an operator can control without maintaining skin in the game. They may take the form of formal bonds, additional operator collateral outside the bond and reserve layers tied to vault health. The table reports all three types, each normalized by the amount of stake managed for validation (Rocket Pool, 2026c).

[8] Fork choice refers to the rule that selects which chain tip validators extend. An entity controlling sufficient voting power can bias this rule to censor transactions or trigger reorganizations.

2.1 Major liquid staking protocols

Table 1 summarizes major liquid staking protocols and reports selected institutional features. We exclude centralized or custodial LSTs such as WBETH (Binance) and cbETH (Coinbase).[9] For each protocol module, the table records the claim-origination type, size, capital structure decomposed into the three layers defined above, whether operator admission is permissioned, the headline reward fee and the denomination of the primary first-loss layer.

Table 1: Major liquid staking protocols: institutional features, capital structure and design dimensions.

			Size		Capital structure					Design	
Protocol	Token	Type	TVL	Share	Bond	Op. ext.	Res.	Total	Perm.	Fee	1st-loss
Lido (Curated)	stETH	Comp.	8.64	24.2	0	0	0	0	Y	10%	n/a
Lido (CSM)					0.075	0	0	0.075	N	10%	stETH
ether.fi	eETH	Comp.	2.71	7.6	0	0	0	0	Y	prot.-set	n/a
Rocket Pool	rETH	Comp.	0.589	1.7	0.25	0.075	0	0.325	N	5–14%	ETH + RPL
StakeWise (Std.)	osETH	Disc.	0.357	1.0	0	0.031	0.10	0.131	N	vault-set	ETH
StakeWise (100%)					0	0.032	0	0.032	Y	≤5%	ETH + SWISE

Notes: Snapshot date 2025-12-13. Type classifies claim origination as compounding (par entry, value through rebasing) or discounting (collateralized minting against a retained slice). Each capital ratio is per unit of ETH managed; a zero means no mandatory requirement is documented. TVL is in millions of ETH; Share is the percentage of total staked ETH. Fee records the headline reward skim or operator commission. 1st-loss records the denomination of the primary mandatory first-loss layer. See Appendix I for detailed variable definitions, data sources and protocol-specific computations.

Where a protocol operates multiple modules with distinct capital requirements, the table reports separate rows with the module label in parentheses, and larger totals correspond to more capital at risk per unit of managed stake.

For Lido, the table covers two modules: the Curated Module and the Community Staking Module (CSM), selected to contrast a permissioned operator set against a permissionless entry path with explicit bond requirements (Lido Finance, 2024). Both are compounding. No mandatory per-validator capital is documented for the Curated Module, while the CSM

[9] In these designs, the issuing entity directly controls the validator keys and retains the associated economic exposure in-house. The control-exposure wedge the paper studies arises from the separation of operational control and economic ownership through a pool, which does not apply to custodial staking.

requires a bond that can be posted in ETH, stETH or wstETH (a non-rebasing wrapped form of stETH) (Lido Finance, 2025a). Both modules charge a 10% reward fee, and TVL and share represent the total across both.

For ether.fi, operator admission is permissioned through whitelisting (ether.fi, 2026a,b). The module is compounding and records zeros across all capital layers because no mandatory per-validator capital is documented. Its reward fee is protocol-set and governance-adjustable (ether.fi, 2026c).

For Rocket Pool, entry is permissionless: node operators contribute bonded ETH alongside pooled ETH and post RPL (Rocket Pool's protocol token) as collateral (Rocket Pool, 2026c,a). The module is compounding, with bonded ETH in the Bond column, RPL collateral in outside-bond capital and the reward fee recorded as a 5–14% operator commission (Rocket Pool, 2026b).

For StakeWise, the table distinguishes two vault tiers. Both are discounting because osETH is originated through collateralized minting against a retained collateral slice. In standard vaults (90% loan-to-value, LTV), first-loss capacity combines operator collateral and the mandatory reserve implied by the LTV cap (StakeWise, 2025). In the 100% LTV tier, the reserve is removed and replaced with insurance funded by SWISE (StakeWise's governance token), and the tier is approved and permissioned by a decentralized autonomous organization (DAO) (StakeWise Governance, 2024). TVL and share represent the total across both tiers.

3 Discount model with an exogenous operator pool

This section develops a simple version of the discount model with an exogenous operator pool. Time has two dates. At $t = 1$, operators originate liquid staking claims and sell them to users in competitive secondary markets. At $t = 2$, a strategic actor that has acquired voting power through the protocol chooses whether to attack. Collateral absorbs slashing losses up to its face value, and any residual loss is socialized through lower LST backing. Competitive

pricing then links the date-1 LST price to expected backing. The section concludes by connecting the pool's loss-allocation problem to the attacker's incentive problem. Section 4 later extends the same logic to endogenous operator participation.

3.1 Environment

We summarize the attack profile by the voting power required for a successful attack, denoted by $S > 0$ and measured in ETH. The attacker operates $M^A = S$ units of staked ETH through the liquid staking protocol, thereby controlling exactly the minimum stake required for a successful attack. Let a denote the attacker's private benefit, also measured in ETH.[10] Slashing severity is $\gamma \geq 0$ per unit of attacking stake, so total slashing equals γS. The parameter $d \in [0, 1]$ captures the confidence or valuation loss: it reduces the value of any ETH that remains exposed through the event, so a unit valued at 1 falls to $1 - d$. Throughout this section, ETH is the numeraire, so all payoffs and values are denominated in ETH. For expositional simplicity, the model normalizes the initial token supply so that one unit of LST corresponds to one unit of backing before any loss event, abstracting from rewards, fees, and rebalancing mechanics.

The protocol is summarized by a bond ratio $x \in [0, 1]$ and a proportional protocol fee $\tau \in [0, 1)$ levied on staking rewards. The per-period staking reward rate per unit of staked ETH is $\lambda > 0$, so the net reward rate is $(1 - \tau)\lambda$. We assume that staking rewards are received only by operators that do not attack the protocol. To obtain voting power S through the liquid staking protocol, the attacker contributes S ETH, receives $(1 - x)S$ LST claims, and leaves the remaining xS locked as staked collateral. In an attack, slashing destroys γS units of ETH from the attacking stake. Losses are first absorbed by the operator's locked collateral, up to xS, and any residual loss is socialized to LST holders through lower pooled backing. Let U

[10] This benefit may arise from direct gains enabled by the attack, such as double-spending, or from trading positions built around the attack. For example, Tzinas and Zindros (2023) illustrate an attack based on shorting the liquid staking token.

denote the uncovered slashing loss borne by the pool, so that

$$U = \gamma M^A - \min\Big(\gamma M^A,\ xM^A\Big) = \max\big\{0,\ (\gamma - x)M^A\big\} = \max\{0,\ (\gamma - x)S\}\,. \tag{1}$$

We assume an exogenous mass of honest operators—defined as operators who never attack the protocol, regardless of economic incentives—supplies pooled stake of $M^O > 0$ units of ETH. Total pooled ETH is then

$$M^L := M^O + M^A = M^O + S, \tag{2}$$

and all ETH contributed to the protocol is staked.

Let L denote the total supply of LST claims issued by the protocol, so that

$$L = (1 - x)M^L. \tag{3}$$

Prior to any loss event, the LST is a pro rata claim on the non-collateral backing, $(1 - x)M^L$, abstracting from rewards and fees. It is therefore fully backed on a per-token basis before an attack.

At $t = 1$, the LST trades competitively in the secondary market at price p (ETH per LST) among a mass of representative users. Market participants hold rational expectations about the probability $q \in [0, 1]$ that an attack occurs at $t = 2$ and the attacker can sell its LST units at price p before choosing whether to misbehave. At $t = 2$, the attacker chooses an action $A \in \{0, 1\}$, where $A = 1$ means "attack." In equilibrium, the anticipated probability q coincides with the probability of attack induced by the attacker's optimal strategy (including mixing).

3.2 Attacker's problem

We begin with the corresponding benchmark under traditional staking, in which the full S ETH remains exposed and locked throughout. In the event of an attack, the attacker bears

both the slashing loss and the confidence loss on the full amount S, and also forfeits the staking rewards λ on that amount. Measuring payoffs net of the initial capital outlay S, the attacker's incremental payoff from attacking under traditional staking is

$$\Delta\Pi_{\text{traditional}} = a + S(1-\gamma)(1-d) - \big(1+\lambda\big)S.$$

Liquid staking changes this payoff structure because the attacker can sell the liquid portion $(1-x)S$ of its stake at $t=1$ and leave only the collateral share xS exposed. In the strategy we consider next, the attacker sells its $(1-x)S$ LST claims at price p. Under the no-attack benchmark with $q=0$, competitive pricing implies $p=1$. If the attacker acts honestly, its payoff is simply the net staking reward:

$$\Pi_{\text{liq}}(0) = (1-\tau)\lambda S.$$

If instead it attacks at $t=2$, it receives the private benefit a and loses γS units of ETH through slashing. Losses are absorbed first by collateral up to xS. The value that remains after slashing is therefore $\max\{0,\, x-\gamma\}S$, and because it stays exposed through the event it is further depreciated by $(1-d)$. The resulting payoff under attack is

$$\Pi_{\text{liq}}(1;p) = a + (1-d)\max\{0,\, x-\gamma\}S + p(1-x)S - S.$$

The attacker's incremental payoff under the liquid-staking strategy is

$$\begin{aligned}\Delta\Pi_{\text{liquid}}(p) &\equiv \Pi_{\text{liq}}(1;p) - \Pi_{\text{liq}}(0)\\ &= a + \Big[(1-d)\max\{0,\, x-\gamma\} + p(1-x) - 1 - (1-\tau)\lambda\Big]S.\end{aligned} \tag{4}$$

Attack is privately optimal whenever $\Delta\Pi_{\text{liquid}}(p) \geq 0$, or equivalently whenever p weakly exceeds a profitability threshold. Let $p_{\text{profitable}}$ be defined by $\Delta\Pi_{\text{liquid}}(p_{\text{profitable}}) = 0$:

$$p_{\text{profitable}} = \left(\frac{1}{1-x}\right)\left[1 + (1-\tau)\lambda - (1-d)\max\{0,\, x-\gamma\} - \frac{a}{S}\right]. \tag{5}$$

This cutoff is also robust to the choice of numeraire. Online Appendix A shows that the same cutoff obtains in a USD-denominated version of the model. The reason is that the USD/ETH exchange-rate scales both within-period payoffs and the no-attack continuation value, so it cancels out up to a constant rescaling associated with discounting.

3.3 The backing value of the LST

So far, the attacker has taken the date-1 LST price p as given. We now characterize in Proposition 1 the post-event backing value of the LST, since under risk-neutral competitive pricing, expected backing determines the price that can support any candidate attack outcome.

Proposition 1 (LST backing under slashing with bond coverage)**.** *If the pool has backing M^L and a pre-event LST supply of $L = (1-x)M^L$, then the post-attack backing value per unit of LST is*

$$b = \min\left\{\frac{(1-x)M^L - U}{L},\ 1\right\} = 1 - \max\{0,\ \gamma - x\}\left(\frac{1}{1-x}\right)\frac{S}{M^L}. \tag{6}$$

Proof. The proof is in Appendix II. Intuitively, the operator collateral is first-loss capital. When $x \geq \gamma$, the collateral fully absorbs the slashing loss and backing remains at par. When $x < \gamma$, the uncovered loss $(\gamma - x)S$ is spread across the non-collateral claim base $(1-x)M^L$, which gives the stated decline in backing per token. ■

3.4 Equilibrium

Let $V(A)$ denote the $t = 2$ backing value of one unit of the LST, measured in ETH, under action $A \in \{0, 1\}$. Under no attack ($A = 0$), this value is equal to 1, while under attack ($A = 1$), it is equal to b. Risk-neutral, competitive LST users price the token at $t = 1$ at its expected backing value, given their information and their correct anticipation of the attacker's behavior. Thus,

$$p = \mathbb{E}[V(A) \mid \mathcal{I}_1].$$

Under a deterministic (pure-strategy) equilibrium, this reduces to $p = V(A)$. Given p, the attacker chooses $A = 1$ if and only if $\Delta\Pi_{\text{liquid}}(p) \geq 0$.

A mixed equilibrium requires the attacker to be indifferent between attacking and not attacking, which is equivalent to

$$p = p_{\text{profitable}}. \tag{7}$$

Lemma 1 (Exogenous-scale equilibrium outcomes)**.** *The following statements hold:*

(i) If $p_{profitable} \geq 1$, then there exists a pure-strategy equilibrium with no attack ($q = 0$) and price $p = 1$.

(ii) If $b \geq p_{profitable}$, then there exists a pure-strategy equilibrium with attack ($q = 1$) and price $p = b$.

(iii) If $1 > p_{profitable} > b$, then no pure-strategy equilibrium exists, and there exists a mixed-strategy equilibrium in which the attacker mixes with attack probability $q \in (0, 1)$, the date-1 price satisfies the indifference condition (7), and $q = \frac{1-p_{profitable}}{1-b}$.

Proof. The proof is in Appendix II. Intuitively, the no-attack and attack candidates are evaluated at the prices implied by the pure outcomes, namely $p = 1$ and $p = b$. If neither boundary price is self-enforcing, beliefs adjust until the indifference condition (7) holds, which yields the mixed case. ■

Lemma 1 can be read as a three-way comparison between $p_{\text{profitable}}$ and the two boundary prices 1 and b. Region (i) is the case in which the no-attack price $p = 1$ already deters attack. Region (ii) is the case in which even the attack price $p = b$ leaves attack privately profitable. Region (iii) is the intermediate case in which neither pure outcome is self-enforcing, so q adjusts until the indifference condition (7) holds.

To restate the equilibrium classification directly in terms of the bond ratio x, note that backing under attack remains at par when $x \geq \gamma$ and falls below one when $x < \gamma$. This distinction gives rise to two cutoff values,

$$x_0 := \frac{a}{S} - (1-\tau)\lambda, \quad x_1 := \frac{\frac{a}{S} - (1-\tau)\lambda - \frac{S}{M^L}\gamma}{1 - \frac{S}{M^L}}.$$

where x_0 is the par-price deterrence cutoff and x_1 is the attack-price indifference cutoff. Proposition 2 then maps the no-attack, attack, and mixed regions into conditions on x, distinguishing the full-coverage region $x \geq \gamma$ from the uncovered-loss region $x < \gamma$.

Proposition 2 (Exogenous-scale equilibrium by bond ratio)**.** *The following statements hold:*

(i) If $x_0 \leq x < \gamma$, or if $x \geq \gamma$ and $(1-\tau)\lambda + dx + (1-d)\gamma \geq \frac{a}{S}$, then there exists a pure-strategy equilibrium with

$$q = 0, p = 1.$$

(ii) If $x \geq \gamma$ and $(1-\tau)\lambda + dx + (1-d)\gamma < \frac{a}{S}$, then there exists a pure-strategy equilibrium with

$$q = 1, p = 1.$$

If $x < \gamma$ and $x < \min\{x_0, x_1\}$, then there exists a pure-strategy equilibrium with

$$q = 1, p = 1 - \left(\frac{S}{M^L}\right)\left(\frac{\gamma - x}{1 - x}\right).$$

(iii) If $x < \gamma$ and $\min\{x_0, x_1\} < x < \max\{x_0, x_1\}$, then there exists a mixed-strategy equilibrium with

$$p = \left(\frac{1}{1-x}\right)\left[1 + (1-\tau)\lambda - \frac{a}{S}\right],$$

and

$$q = \frac{\frac{a}{S} - (1-\tau)\lambda - x}{\frac{S}{M^L}(\gamma - x)} \in (0, 1).$$

Proof. The proof is in Appendix II. Intuitively, the proposition translates Lemma 1 into bond-ratio terms. Substituting the x-dependence of $p_{\text{profitable}}$ and b turns each regime boundary into a cutoff inequality. The objects x_0 and x_1 are the bond-ratio values at which those inequalities change sign. ■

3.5 Security implications of liquid staking

The model yields a simple comparison of deterrence with and without liquid staking for a given attack profile (d, S, γ). Here we study how small the private benefit a must be for the attacker to refrain from attacking. In the traditional-staking benchmark, the full S ETH remains exposed throughout. Deterrence is therefore characterized by $\Delta\Pi_{\text{traditional}} \leq 0$, or equivalently

$$a \leq a^*_{\text{traditional}} := S\Big(\lambda + d + (1-d)\gamma\Big). \tag{8}$$

With liquid staking, the attacker can separate control from exposure by obtaining voting power S using pooled stake and then selling the liquid claim before choosing an action. The attacker can sell only $(1-x)S$ LST claims at date 1. The remaining xS stays exposed as collateral. Any uncovered slashing loss is passed through to the pool and therefore to whoever holds the LST after the sale. Proposition 2 then implies that the no-attack outcome with $p = 1$ is sustainable if and only if attack is not privately profitable at $p = 1$, that is, if and only if $p_{\text{profitable}} \geq 1$. Using equation (5), this condition is equivalent to

$$a \leq a^*_{\text{liquid}}(x) := S\Big((1-\tau)\lambda + dx + (1-d)\min(\gamma, x)\Big). \tag{9}$$

Comparing equations (8) and (9) immediately yields the next corollary on how liquid staking changes deterrence.

Corollary 1 (Liquid staking weakens deterrence under limited bonding)**.** *If $x \leq 1$, then the deterrence threshold under the liquid-staking strategy in equation (9) is weakly smaller than the traditional-staking threshold in equation (8):*

$$a^*_{liquid}(x) \leq a^*_{traditional}.$$

Proof. The proof is in Appendix II. Intuitively, liquid staking allows the attacker to offload part of the position before attacking. As a result, the portion of value that remains exposed to slashing and depreciation is weakly smaller than under traditional staking. The deterrence

cutoff therefore cannot exceed its traditional-staking counterpart. ■

When $d = 0$, the weaker deterrence under liquid staking does not by itself generate any additional inefficiency from the socialization of slashing losses. In the region where $x < \gamma$, the attacker's private profitability threshold is independent of γ. Socialization then affects equilibrium attack probabilities only through prices, without altering the ETH-denominated surplus of either the attacker or secondary-market buyers.

The conclusion changes when $d > 0$. Lower private exposure of liquid staking operators increases the equilibrium attack probability and, with it, the expected incidence of post-attack ETH depreciation. Because p is denominated in ETH per LST, it internalizes expected backing losses within the pool, but not a decline in ETH's external value. The resulting welfare loss is therefore borne by all ETH holders, whether they participate in liquid staking or not. In this sense, the depreciation channel is the source of inefficiency in this frictionless benchmark.

Lemma 2 (Par-pricing benchmark (p exogenously fixed at one))**.** *Consider a benchmark in which the secondary-market price at $t = 1$ is exogenously fixed at $p = 1$ and does not respond to information about attack incentives. Any mixed equilibrium with $q \in (0, 1)$ is a knife-edge case requiring $p_{profitable} = 1$. Hence, in general, the mixed region disappears. A pure-strategy no-attack equilibrium exists if and only if $p_{profitable} \geq 1$. If instead $p_{profitable} < 1$, the unique pure-strategy equilibrium features attack.*

Proof. With p fixed at 1, any mixed equilibrium requires the indifference condition (7), hence $p_{\text{profitable}} = 1$. This is a knife-edge case, so in general the mixed region disappears. The pure-strategy claims then follow because the attacker does not deviate from no attack if and only if $\Delta\Pi_{\text{liquid}}(1) \leq 0$, which by the definition of $p_{\text{profitable}}$ is equivalent to $p_{\text{profitable}} \geq 1$. If $p_{\text{profitable}} < 1$, then $\Delta\Pi_{\text{liquid}}(1) > 0$, so the attacker strictly prefers attack and the unique pure-strategy equilibrium features attack. ■

Lemma 2 isolates the role of price adjustment. The no-attack region is unchanged: whenever $p_{\text{profitable}} \geq 1$, the no-attack outcome with $p = 1$ is an equilibrium under endogenous pricing and remains one under par pricing. What disappears is the mixed region

$b < p_{\text{profitable}} < 1$. Under endogenous pricing, the market can lower the price until the indifference condition (7) holds, yielding a mixed equilibrium with $q = \frac{1-p_{\text{profitable}}}{1-b} \in (0,1)$. Under par pricing, price cannot fall enough for the indifference condition (7) to hold, so the mixed equilibrium disappears. Since $p_{\text{profitable}} < 1$ implies that attack is strictly profitable at $p = 1$, the outcome collapses to the pure-attack action ($q = 1$) throughout that region.

3.6 Illustrations

Figure 1 illustrates the equilibrium of the exogenous participation model under the baseline calibration $M^L = 1$, $S = 2/3$, $\gamma = 0.45$, $a = 0.25$, and $d = 0.4$, with $(\lambda, \tau) = (0.095, 0.5)$ and $x \in [0, 0.75]$. Under this calibration, attack would not occur under traditional staking, since $\Delta\Pi_{\text{traditional}} = a - S\big(\lambda + d + (1-d)\gamma\big) = 0.25 - (2/3)\big(0.095 + 0.4 + 0.6 \cdot 0.45\big) = -0.26 < 0$. The figure therefore highlights how liquid staking changes the equilibrium regime and, as summarized in Proposition 2, shows that a higher collateral requirement increases the attacker's exposure and shrinks the mixed-attack region.

Panel (a) of Figure 1 plots the deterrence thresholds $a^*_{\text{traditional}}$ and $a^*_{\text{liquid}}(x)$ from Section 3.5, that is, the maximum private benefit consistent with no attack. Under traditional staking, $a^*_{\text{traditional}} = S\big(\lambda + d + (1-d)\gamma\big)$ is constant in x because the attacker must keep the full S ETH exposed through the event. Under liquid staking, $a^*_{\text{liquid}}(x)$ increases with x because a higher collateral requirement leaves more value exposed to depreciation and slashing. As x approaches one, the liquid-staking threshold approaches the traditional one, but it remains slightly below it because liquid staking modifies traditional staking's opportunity-cost term λ by adjusting it for the protocol fees, $(1-\tau)\lambda$.

Panel (b) compares the two objects that determine equilibrium under frictionless pricing. The curve $b(x)$ is the LST backing value and therefore the price consistent with attack, while $p_{\text{profitable}}(x)$ is the minimum date-1 price at which attack remains privately profitable. A pure-attack equilibrium can therefore be sustained only when the price implied by attack is high enough to satisfy the attacker's incentive constraint, that is, when $b \geq p_{\text{profitable}}$. As collateral coverage increases, b rises toward one because a larger collateral buffer reduces

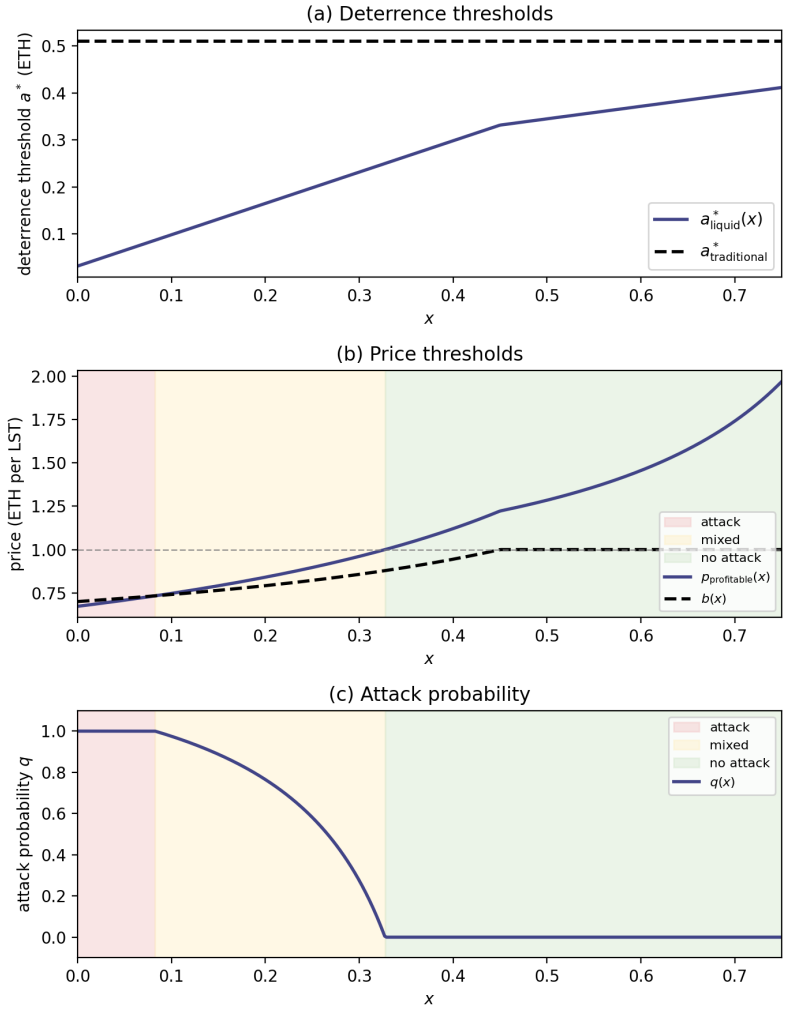

Figure 1: Equilibrium outcomes under frictionless pricing for varying bond ratios x. Panel (a) plots $a^*_{\text{traditional}}$ and $a^*_{\text{liquid}}(x)$. Panel (b) plots $p_{\text{profitable}}(x)$ and $b(x)$. Panel (c) plots $q(x)$.

the uncovered loss borne by honest LST holders. By contrast, $p_{\text{profitable}}$ reflects the attacker's private cost of misbehavior through retained collateral exposure and the opportunity-cost term $(1-\tau)\lambda$. The relative position of the two curves then determines whether the equilibrium features no attack, pure attack, or mixing in the intermediate region $1 > p_{\text{profitable}} > b$, as summarized in Proposition 2.

Panel (c) maps the pricing comparison in panel (b) into equilibrium attack probabilities. When $p_{\text{profitable}} \geq 1$, the no-attack price $p = 1$ already deters attack, so $q = 0$. When $b \geq p_{\text{profitable}}$, the price consistent with attack still makes attack privately profitable, so $q = 1$. In the intermediate region $1 > p_{\text{profitable}} > b$, neither pure outcome is self-enforcing, so the attacker mixes. The indifference condition (7) then combines with competitive pricing $p = (1-q)\cdot 1 + q \cdot b$ to yield the attack probability $q = (1 - p_{\text{profitable}})/(1-b)$. As x rises, the attacker retains more collateral exposure, which shrinks the mixed region and lowers q wherever mixing occurs.

4 Discount model with an endogenous operator pool

We now extend the discount model by endogenizing operator participation. Honest stake M^O becomes the key scale variable, and it is jointly determined with the date-1 LST price p and the rational-expectations attack probability q.

4.1 Environment

As in Section 3, ETH pooled into the liquid-staking protocol and total LST supply are given by equations (2) and (3), respectively. Claims backed by attacker deposits are sold at $t = 1$ to secondary-market buyers.

We now assume that users require a net return of $\mu \geq 0$ per unit of ETH allocated to LSTs, which captures the return on their outside option.[11] This implies an additional discount

[11] In principle, the outside option for both users and operators is traditional staking, which yields a gross return of $1 + \lambda$. We capture the practical advantages of liquid staking, such as liquidity, lower delegation costs and DeFi integration, in reduced form by parameterizing the outside-option return as $\mu \geq 0$. The case $\mu < \lambda$

factor of $(1+\mu)^{-1}$ in the asset's expected return. Using the notation from Section 3, the participation condition becomes

$$p = \left(\frac{1}{1+\mu}\right)\left[(1-q)\cdot 1 + q\cdot b\right]. \tag{10}$$

Online Appendix B studies traditional staking as a no-attack outside option. It shows that this additional participation margin may eliminate some positive-attack equilibria, but it does not change the optimality of a policy that already implements a no-attack equilibrium in the design problem studied below.

Let M denote total ETH supply and write $M \equiv M^U + M^S$, where M^U is unstaked ETH and M^S is total staked ETH. We decompose total staked ETH as

$$M^S \equiv M^T + M^L,$$

where M^T denotes ETH staked through traditional staking and held by honest operators. Traditional staking accounts for a constant share α of total ETH supply, so that $M^T = \alpha M$. After normalizing $M = 1$, this implies $M^T = \alpha$.

We also endogenize the ETH staking return by adopting Ethereum's protocol minting rule, under which the issuance rate declines with total staked ETH. For simplicity, we assume that rewards come exclusively from issuance,[12] so that

$$\lambda = \frac{k}{\sqrt{M^S}}, \tag{11}$$

where $k > 0$ is a constant. The next subsection states the honest-operator participation condition. The attacker's problem then links M^S to the required attack stake S and the decomposition (M^O, M^A) in equilibrium.

---

then corresponds to a net advantage of liquid staking.

[12] In practice, part of operators' compensation comes from priority fees voluntarily paid by users who want their transactions processed first. We abstract from this source of revenue because it does not affect the qualitative results of the model.

4.2 Participation condition of honest operators

Since total operator profit equals per-unit profit times operated stake, the expected profit of honest operators is

$$\begin{aligned}\Pi^O &= (1-q)\Big[x + (1-\tau)\lambda + p(1-x) - 1\Big]M^O \\ &\quad + q\Big[(1-d)\big(x + (1-\tau)\lambda\big) + p(1-x) - 1\Big]M^O \\ &= \Big[(1-\tau)\lambda - (1-p)(1-x) - qd\big[x + (1-\tau)\lambda\big]\Big]M^O.\end{aligned} \tag{12}$$

The first line of equation (12) expresses expected profit as the weighted average of the no-attack payoff and the attack payoff, under the assumption that honest operators sell their LST claims in the secondary market at $t = 1$ regardless of their preferred strategy. Under attack, the exposed continuation value, $xM^O + (1-\tau)\lambda M^O$, is reduced by the factor $(1-d)$.

We assume that honest operators also have an outside option with net return μ. Measured in ETH, the relevant capital base consists of the locked collateral, xM^O, plus the portion of the remaining capital lost to the secondary-market discount, $(1-p)(1-x)M^O$. The participation condition is

$$\Pi^O \geq \mu[x + (1-p)(1-x)]M^O, \tag{13}$$

with equality in equilibrium.

Lemma 3 (Honest-operator participation condition)**.** *In equilibrium, honest-operator participation satisfies*

$$[(1-\tau)\lambda - (1-p)(1-x) - qd[x + (1-\tau)\lambda]]\, M^O = \mu[x + (1-p)(1-x)]M^O. \tag{14}$$

Proof. Follows immediately from equations (12) and (13) in equilibrium. ■

4.3 Attacker's problem

This subsection carries over the liquid-staking attack strategy from Section 3, but now honest-operator participation jointly determines total staking and the required attack stake.

We assume that an attack requires control of a fraction $f \in (0, 1)$ of total staked ETH:

$$S = fM^S = f(M^T + M^O + M^A),$$

which the attacker acquires through the liquid staking protocol, so $M^A = S$. Substituting into the stake requirement gives

$$M^S = \left(\frac{1}{1-f}\right)\left[\alpha + M^O\right], \qquad S = \left(\frac{f}{1-f}\right)\left[\alpha + M^O\right]. \tag{15}$$

Pooled stake in the liquid staking protocol is therefore

$$M^L = M^O + M^A = \left(\frac{1}{1-f}\right)\left[f\alpha + M^O\right]. \tag{16}$$

For the attacker to obtain the required stake through the liquid staking protocol, total liquid-staked ETH must satisfy $M^L \geq S$. Under equation (16), this condition is equivalent to $M^O \geq 0$. Finally, feasibility requires that liquid staking not exceed the ETH supply outside traditional staking, that is, $M^L \leq 1 - \alpha$. We impose this feasibility condition throughout.

4.4 Equilibrium

For a fixed bonding policy $x \in [0, 1]$, the equilibrium objects are the attack probability $q \in [0, 1]$, the honest stake $M^O \geq 0$ and the secondary-market price p. Given M^O, equations (15) and (16) pin down M^S, S and M^L. Equations (5), (6) and (10) then determine $p_{\text{profitable}}$, b and p. In interior cases, equilibrium also requires the indifference condition (7). We call an equilibrium operating when $M^O > 0$ and shutdown when $M^O = 0$. The proposition below classifies no-attack, attack, mixed and shutdown outcomes. The mixed case differs

across $x < \gamma$ and $x \geq \gamma$ because only the bond-limited region has backing below par $b < 1$ in the event of attack.

Proposition 3 (Endogenous-scale equilibrium classification)**.** *Fix a bond ratio $x \in [0, 1]$. The following statements hold:*

(i) A no-attack operating equilibrium exists if and only if some $M^O > 0$ satisfies equation (14) at $q = 0$ and

$$p_{profitable} \geq \frac{1}{1+\mu},$$

in which case

$$q = 0, \qquad p = \frac{1}{1+\mu}.$$

(ii) An attack operating equilibrium exists if and only if some $M^O > 0$ satisfies equation (14) at $q = 1$ and

$$b \geq (1+\mu)p_{profitable},$$

in which case

$$q = 1, \qquad p = \frac{1}{1+\mu}b.$$

(iii) A mixed operating equilibrium exists if and only if there is a pair (q, M^O) with $q \in (0, 1)$ and $M^O > 0$ satisfying equation (14) and the indifference condition (7). If $x < \gamma$, then

$$q = \frac{1-(1+\mu)p_{profitable}}{1-b} \in (0, 1).$$

If $x \geq \gamma$, then a mixed equilibrium requires

$$p_{profitable} = \frac{1}{1+\mu}, \qquad q \in (0, 1).$$

(iv) If no pair (q, M^O) satisfies (i), (ii) or (iii), then the equilibrium is shutdown with $M^O = 0$.

Proof. See Appendix II for the proof. Intuitively, the no-attack and attack cases evaluate the attacker's incentive at the prices implied by the boundary regimes. In mixed cases, the indifference condition (7) replaces the boundary inequalities. Under full coverage ($x \geq \gamma$), the mixed condition collapses to the knife-edge equality $p_{\text{profitable}} = \frac{1}{1+\mu}$. ■

A sufficient condition for shutdown is that μ exceed the maximum feasible net staking return. Since λ decreases with aggregate stake, it reaches its maximum at $M^O = 0$. The corresponding net return is $(1-\tau)\lambda(0) = (1-\tau)k\sqrt{1-f}/\sqrt{\alpha}$. If $\mu > (1-\tau)\lambda(0)$, the operator participation condition $(1-\tau)\lambda = \mu$ has no solution with $M^O > 0$. In that case, none of Cases (i) to (iii) admits an operating equilibrium and the outcome is shutdown.

4.5 Security implications of liquid staking

We now compare the deterrence cutoffs in the private benefit a when protocol scale is endogenous. For a fixed bonding policy $x \in [0,1]$, consider a candidate no-attack equilibrium with $q = 0$ and price $p = \frac{1}{1+\mu}$. As before, let $a^*_{\text{traditional}}$ denote the deterrence cutoff under traditional staking, and let $a^*_{\text{liquid}}(x)$ denote the corresponding cutoff under liquid staking. Corollary 2 shows that the exogenous ranking extends to the endogenous model under a simple sufficient condition. When collateral coverage is strong or the depreciation component of attack losses is large, liquid staking cannot reduce the attacker's retained exposure by much, so its deterrence cutoff remains weakly below the traditional benchmark.

Corollary 2 (Liquid staking weakens deterrence under limited bonding)**.** *Suppose $\mu > 0$, there exists some $M^O > 0$ satisfying honest-operator participation equation (14) at $q = 0$, and*

$$\frac{\tau}{1-\tau}\mu + \frac{1-\gamma}{1+\mu} \geq (1-\gamma)(1-d).$$

Then the endogenous deterrence cutoff for liquid staking is weakly smaller than the corresponding cutoff for traditional staking for all $x \in [0,1]$:

$$a^*_{liquid}(x) \leq a^*_{traditional}.$$

Proof. See Appendix II for the proof. Intuitively, once the common no-attack scale is fixed, endogenous participation introduces an additional discount through the no-attack price $p = \frac{1}{1+\mu}$. That discount partly offsets the attacker's ability to unload the liquid portion of the stake. When collateral coverage is sufficiently strong or the depreciation component is sufficiently large, the remaining offloading advantage is too small to overturn the traditional ranking. ■

Risk-neutral demand internalizes expected uncovered slashing into the LST price through equation (10), but that price response does not restore deterrence. Liquid exit and limited collateral still reduce the value the attacker keeps exposed through the event. Endogenous scale adds a participation channel. In the bond-limited region ($x < \gamma$), attack risk and backing dilution affect p and therefore honest-operator participation equation (14) through the base-capital term $[x + (1-p)(1-x)]M^O$. When $x \geq \gamma$, we have $b = 1$, so equation (10) implies $p = \frac{1}{1+\mu}$ for all q. Any interior mixing then requires the indifference condition (7) to hold, which here reduces to the equality $p_{\text{profitable}} = \frac{1}{1+\mu}$, together with honest-operator participation equation (14).

Lemma 4 (Par-pricing benchmark with endogenous scale)**.** *Consider a benchmark in which the secondary-market price at $t = 1$ is exogenously fixed at the no-attack level $p = \frac{1}{1+\mu}$ and does not respond to information about attack incentives. Any mixed operating equilibrium with $q \in (0,1)$ is a knife-edge case requiring $p_{profitable} = \frac{1}{1+\mu}$. Hence, in general, the mixed region disappears. A no-attack operating equilibrium exists if and only if there exists some $M^O > 0$ satisfying honest-operator participation (14) at $q = 0$ and $p_{profitable} \geq \frac{1}{1+\mu}$. If instead there exists some $M^O > 0$ satisfying honest-operator participation (14) at $q = 1$ and $p_{profitable} < \frac{1}{1+\mu}$, then an operating equilibrium features attack.*

Proof. With p fixed at $\frac{1}{1+\mu}$, any mixed operating equilibrium requires the indifference condition (7), hence $p_{\text{profitable}} = \frac{1}{1+\mu}$. This is a knife-edge case, so in general the mixed region disappears. The boundary claims then follow immediately: a no-attack candidate requires $p_{\text{profitable}} \geq \frac{1}{1+\mu}$ and an attack candidate requires $p_{\text{profitable}} < \frac{1}{1+\mu}$, with some $M^O > 0$ satisfying honest-operator participation equation (14) at $q = 0$ or $q = 1$, respectively. ■

Lemma 4 isolates the role of price adjustment in the endogenous-scale model. Under risk-neutral pricing, the market can lower p until the indifference condition (7) holds and thereby support interior mixing. With par pricing, that adjustment is unavailable, so the mixed region collapses to the knife-edge equality $p_{\text{profitable}} = \frac{1}{1+\mu}$. The remaining operating candidates are the two boundary regimes. If participation fails at both boundaries, the outcome is shutdown.

4.6 Equilibrium properties

This subsection studies how equilibrium honest stake varies across operating branches. Lemma 5 characterizes the no-attack benchmark scale M_0^O, and Lemma 6 compares any positive-attack operating equilibrium to that benchmark. Proposition 4 then shows that, along a positive-attack branch, the effect of the bond ratio x on honest stake is generally ambiguous.

Lemma 5 (No-attack scale and bond-ratio independence)**.** *Suppose $\mu > 0$ and there exists an operating equilibrium with $q = 0$ and $M^O > 0$. Then M^O satisfies*

$$(1-\tau)\lambda = \mu,$$

and therefore M^O is independent of x. Moreover, M^O is uniquely pinned down by

$$M_0^O = \left(\frac{(1-\tau)k\sqrt{1-f}}{\mu} \right)^2 - \alpha.$$

Proof. If $q = 0$, then equation (10) implies $p = \frac{1}{1+\mu}$. Substituting into the honest-operator participation condition (14) yields

$$(1-\tau)\lambda - (1-p)(1+\mu)(1-x) = \mu x.$$

Using $(1-p)(1+\mu)=\mu$ gives $(1-\tau)\lambda=\mu$, which pins down M^O through $\lambda=\frac{k\sqrt{1-f}}{\sqrt{\alpha+M^O}}$ and is independent of x. ■

Lemma 6 (Positive attack risk lowers scale)**.** *Suppose there exists an operating equilibrium with $q>0$ and $M^O>0$. Then $M^O \le M_0^O$, with strict inequality if $d>0$ or $x<\gamma$.*

Proof. See Appendix II for the proof. Intuitively, positive attack risk lowers operator surplus through expected depreciation and, when $x<\gamma$, through lower backing and price. To restore the break-even point, equilibrium scale must shrink, which raises staking returns. ■

Lemma 6 compares any positive-attack operating equilibrium to the no-attack benchmark. Positive attack risk lowers operators' expected net return through expected depreciation and, when $x<\gamma$, through lower backing and therefore a lower price. Because λ is decreasing in aggregate stake, restoring participation requires a smaller equilibrium scale M^O. The only exception is the knife-edge case with $d=0$ and $x\ge\gamma$ simultaneously, where both channels are absent. The next proposition examines the comparative statics of x along a positive-attack branch.

Proposition 4 (Bonding and honest scale under positive attack risk)**.** *Suppose there exists an operating equilibrium with $q>0$ and $M^O>0$. Then the honest-operator participation condition (14) can be written as*

$$(1-\tau)\lambda-\frac{\mu}{1-qd}-\frac{qdx}{1-qd}-\frac{q(1-x)\big(1-b\big)}{1-qd}=0.$$

As x rises, three channels matter:

1. *Socialized slashing:*
$$-\frac{q(1-x)\big(1-b\big)}{1-qd}.$$
It tends to increase equilibrium honest stake M^O.

2. *Depreciation exposure:*
$$-\frac{qdx}{1-qd}.$$

It tends to decrease equilibrium honest stake M^O.

3. *Discount deterrence: when $x < \gamma$, it operates through equilibrium attack risk q and can also decrease equilibrium honest stake M^O.*

Accordingly, along positive-attack equilibria the effect of x on equilibrium honest stake M^O is not signed in general.

Proof. See Appendix II for the proof. Intuitively, higher bonding x reduces socialized slashing on the liquid portion of stake, but it also increases the value left exposed to depreciation and changes equilibrium attack incentives through q. Because these effects work in opposite directions, the effect on honest scale M^O is not signed. ■

These comparative statics are branch-specific, so they do not imply that a given bond ratio selects a unique equilibrium. As Proposition 3 shows, the endogenous-scale model can admit multiple self-consistent pairs (q, M^O) for the same x. A high-entry branch has higher M^O and therefore a higher attack threshold S. A low-entry branch has lower M^O and lower S, which can sustain higher attack risk because attacker profitability depends on the private benefit per unit of required attack stake. This feedback between operator participation and the attack threshold motivates the protocol-design problem in the next subsection.

4.7 The protocol's problem

We now turn from positive analysis to policy design. At this stage, the protocol chooses the bonding requirement x and the fee share τ, taking into account how these instruments affect equilibrium through operator participation and the attacker's incentives. The key tradeoff is that bonding can deter attacks, but it can also distort operator participation and reduce equilibrium scale, which in turn affects fee revenue.

The protocol's payoff is fee revenue from operated stake. With fee share τ and staking reward λ, the protocol earns $\tau\lambda$ per unit of pooled stake that behaves honestly. Because total

pooled stake M^L includes both honest and attacking stake, protocol fee revenue is

$$\Pi^P := \tau\lambda\left(M^O + \mathbf{1}_{\{A=0\}}M^A\right), \tag{17}$$

where $\mathbf{1}_{\{A=0\}}$ is an indicator for the no-attack state. Lemma 7 shows that this payoff is strictly increasing in M^O.

Lemma 7 (Protocol payoff and honest stake)**.** *For $\tau \in (0, 1]$, Π^P is strictly increasing in $M^O \geq 0$.*

Proof. See Appendix II for the proof. Intuitively, a larger honest stake increases the amount of stake on which the protocol earns fees. Although the reward rate falls with scale, total fee revenue still rises with M^O. ■

For a fixed policy (x, τ) with $x \in [0, 1]$ and $\tau \in (0, 1]$, the following result holds.

Proposition 5 (LST Protocol prefers the no-attack regime)**.** *If an operating equilibrium with $q = 0$ exists, then any operating equilibrium with $q > 0$ yields weakly lower protocol payoff, with strict inequality if $d > 0$ or $x < \gamma$.*

Proof. By Lemma 5, any operating equilibrium with $q = 0$ has honest stake $M^O = M_0^O$, independently of x. By Lemma 6, any operating equilibrium with $q > 0$ has $M^O \leq M_0^O$, with strict inequality when $d > 0$ or $x < \gamma$. Lemma 7 then implies that protocol fee revenue is strictly increasing in M^O. Any operating equilibrium with $q > 0$ therefore yields weakly lower Π^P than the operating equilibrium with $q = 0$, with strict inequality when $d > 0$ or $x < \gamma$. ■

The result in Proposition 5 is consistent with liquid staking practice, where protocols rely on curation, permissioning or bonding to support no-attack outcomes.

Let $M_0^O > 0$ denote the equilibrium honest stake in a no-attack operating equilibrium, as characterized in Lemma 5. The associated attack threshold is

$$S_0 := \left(\frac{f}{1-f}\right)\left[\alpha + M_0^O\right].$$

This expression shows that the attack stake is increasing in the scale of honest operator participation. Proposition 6 combines this expression with the no-attack condition $p_{\text{profitable}} \geq \frac{1}{1+\mu}$ from Proposition 3(i) to characterize the bond ratio x that implements the no-attack regime.

Proposition 6 (Attack feasibility and the bond cutoff at $q = 0$)**.** *At the $q = 0$ operating candidate $M^O = M_0^O$, if $x \leq \gamma$ then the no-attack condition $p_{profitable} \geq \frac{1}{1+\mu}$ is*

$$x \geq 1 - (1+\mu)^2 + \frac{(1+\mu)a}{\left(\frac{f}{1-f}\right)\left[\alpha + M_0^O\right]}. \tag{18}$$

If $x > \gamma$, the same condition is equivalent to

$$\frac{1}{1-x}\left[1 + \mu - (1-d)(x-\gamma) - \frac{a}{\left(\frac{f}{1-f}\right)\left[\alpha + M_0^O\right]}\right] \geq \frac{1}{1+\mu}.$$

Proof. See Appendix II for the proof. Intuitively, the cutoff follows by evaluating the no-attack condition $p_{\text{profitable}} \geq \frac{1}{1+\mu}$ at the $q = 0$ candidate, subject to the operating constraint $M^L \geq S$. ■

Thus, implementation becomes easier when the no-attack benchmark has a larger equilibrium honest stake M_0^O. Since S_0 increases with M_0^O, the right-hand side of equation (18) decreases as M_0^O rises. The minimum bond ratio x needed to sustain the no-attack regime therefore decreases as well.

The protocol's problem is therefore one of implementation rather than preference. Bonding can shrink the set of attack-feasible equilibria, but it need not select a unique no-attack branch because the endogenous-scale model can admit multiple self-consistent equilibria. Even when $x \geq \gamma$ removes uncovered-loss pricing and generic interior mixing, a low-entry branch with lower M^O and a lower attack threshold S may still sustain a boundary attack equilibrium.

This limitation suggests that robust implementation of the no-attack regime typically

requires more than bonding. Some tools act directly on the operator set, including permissioned participation, screening, due diligence, delegated-stake caps and reputation-based admission rules. Others raise effective attacker exposure or reduce the private benefit from attack, for example insurance posted outside the protocol, legal or contractual recourse, slashing penalties that extend beyond the on-chain bond and governance rules that speed recovery after an attack. A third group weakens the uncovered-loss and participation channels through reserve funds, insurance pools, other loss-absorption arrangements that reduce effective socialized slashing and fee policies that improve the return to honest participation. In the current model, permissioned participation is the most direct instrument because it restricts the set of agents who can obtain operational control.

The remaining question is empirical. The protocol can estimate the key primitives governing deterrence and participation, namely a, d, k, α, f, τ and μ, and then assess whether these values place it in a region where the bond requirement delivers a unique no-attack equilibrium. If not, the protocol must rely on complementary tools.

4.8 Illustrations

The previous subsection characterized the design problem analytically. Figures 2 and 3 illustrate how the equilibrium set changes as the bond requirement x varies.

Figure 2 plots, for several values of x, the attacker's indifference locus (7), together with the honest-operator participation locus (14). Candidate operating equilibria are their intersections. In the left panel, raising x moves the system from attack to mixing and then toward no attack, but some intermediate bond ratios still admit more than one intersection. By contrast, the right panel shows that multiplicity can survive even when $x > \gamma$. Once full coverage is reached, uncovered-loss pricing is no longer the source of multiplicity because of full backing, $b = 1$. The coexistence instead comes from scale: a low-entry branch implies a lower attack threshold, so positive attack risk can survive alongside a high-entry no-attack branch.

Figure 3 plots the corresponding equilibrium values of q, M^O, and p as functions of

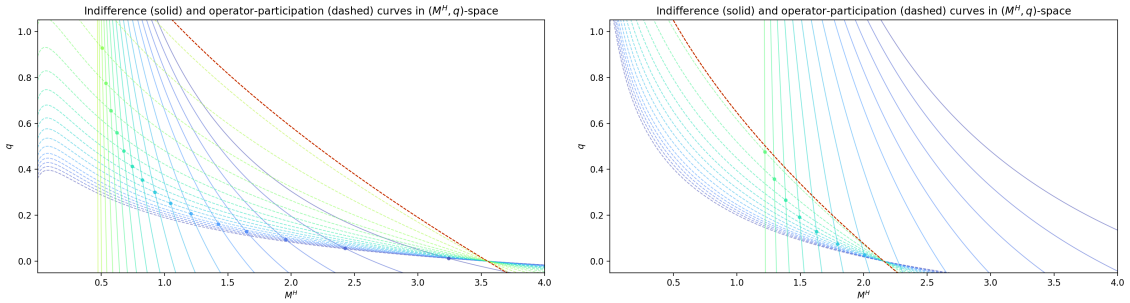

Figure 2: Endogenous-scale curve families as x varies. Each panel plots the attacker's indifference locus (7) and the honest-operator participation locus (14) in the (M^O, q)-space for multiple values of x, with color indicating x and intersections marked. Left panel: baseline calibration with multiplicity concentrated in the bond-limited region, at $f = 0.4$, $\alpha = 0.2$, $k = 0.2$, $\mu = 0.04$, $\tau = 0.5$, $d = 0.4$, $\gamma = 0.30$ and $a = 0.1875$. Right panel: alternative calibration with multiplicity extending into the full-coverage region, at $f = 0.3788$, $\alpha = 0.0215$, $k = 0.1538$, $\mu = 0.0548$, $\tau = 0.3325$, $d = 0.4868$, $\gamma = 0.2251$ and $a = 0.2532$.

the bond ratio x. In the baseline calibration on the left, higher bonding tends to shift the equilibrium set toward lower attack risk and higher honest stake, but not along a single monotone path: the same bond ratio can still support both a no-attack branch and a positive-attack branch. The right panel shows that multiplicity is not confined to the uncovered-loss region. Once $x > \gamma$, price is pinned at $p^*(x) = \frac{1}{1+\mu}$, yet multiple operating branches can still coexist because they differ in scale and therefore in the attack threshold they sustain.

These figures also help interpret the par-pricing benchmark in Lemma 4. Under risk-neutral LST pricing, the mixed branches visible in the figures rely on the date-1 price adjusting until the indifference condition (7) holds. If instead $p = \frac{1}{1+\mu}$ is fixed, that adjustment disappears and the mixed region collapses to the knife-edge case $p_{\text{profitable}} = \frac{1}{1+\mu}$. The remaining operating candidates are the boundary regimes.

5 Conclusion

Liquid staking weakens deterrence because it separates control from loss-bearing. An operator can obtain voting power through pooled stake, sell the liquid claim before the

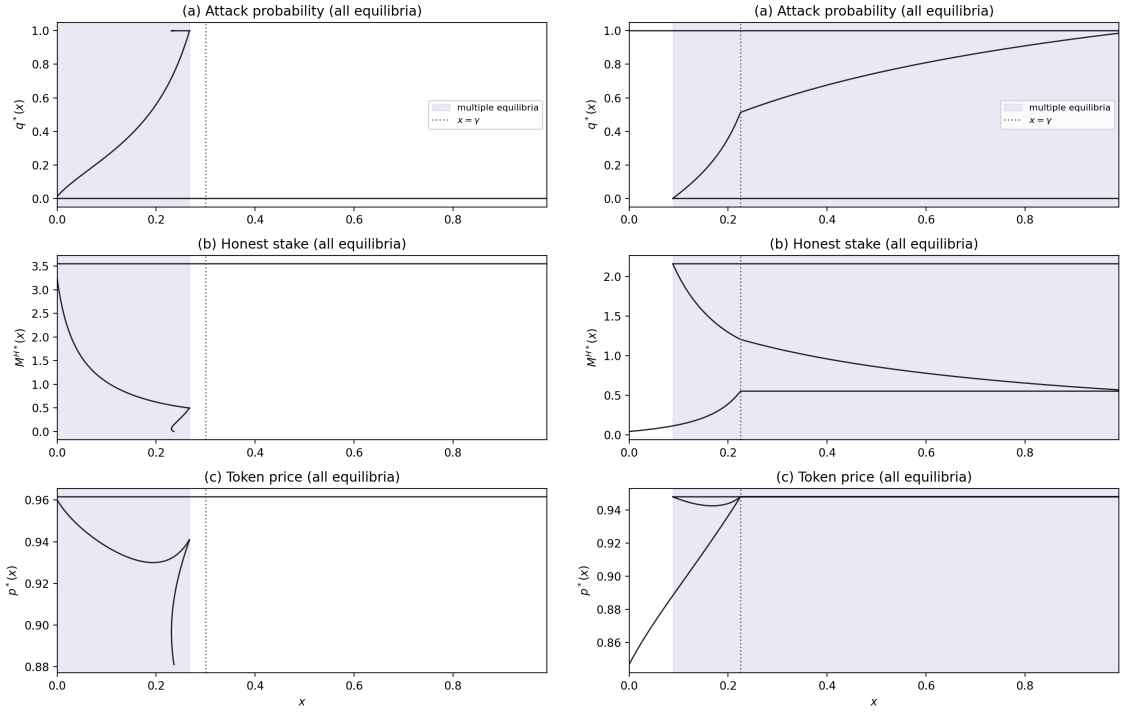

Figure 3: Equilibrium outcomes in the endogenous-scale model as a function of x. In each panel, subplot (a) plots the equilibrium attack probability $q^*(x)$, subplot (b) plots the equilibrium honest stake $M^{O*}(x)$ and subplot (c) plots the equilibrium token price $p^*(x)$. Shaded regions mark bond ratios at which more than one operating equilibrium exists. Left panel: baseline calibration with multiplicity concentrated in the $x < \gamma$ region, at $f = 0.4$, $\alpha = 0.2$, $k = 0.2$, $\mu = 0.04$, $\tau = 0.5$, $d = 0.4$, $\gamma = 0.30$ and $a = 0.1875$. Right panel: alternative calibration with multiplicity for $x > \gamma$, at $f = 0.3788$, $\alpha = 0.0215$, $k = 0.1538$, $\mu = 0.0548$, $\tau = 0.3325$, $d = 0.4868$, $\gamma = 0.2251$ and $a = 0.2532$.

attack decision, and retain only limited exposure to slashing and price declines. Competitive pricing partly offsets this wedge by lowering the LST price when attack risk rises, but it does not eliminate it or internalize the ETH-wide losses borne by native holders.

When scale is endogenous, security and participation are jointly determined. The no-attack benchmark corresponds to the stake level at which honest operators break even. Once attack risk becomes positive, honest stake declines. In this setting, bonding has no monotone effect on scale. On the one hand, it reduces socialized slashing; on the other, it leaves more value exposed to depreciation and, when coverage is incomplete, weakens the price-based deterrence channel. As a result, higher bonding need not increase honest stake or select a unique equilibrium.

The main policy issue is implementation rather than preference. A fee-maximizing protocol may prefer the no-attack outcome whenever it is feasible, because attacks reduce scale and therefore fee revenue. But that does not mean bonding alone can make the no-attack outcome unique. Multiple equilibrium branches can survive at the same bond requirement, so protocols may need additional tools such as curated participation, reserve mechanisms or other institutional safeguards.

The model is deliberately stylized. We leave market frictions, richer strategic interaction among attackers and operators and institutional features such as governance, reputation and delegated selection to future research. Studying them may help clarify how the mechanisms identified here play out in richer market and institutional environments.

References

Beaconcha, "Network Charts, Staked Ether," https://beaconcha.in/charts/staked_ether 2025. Accessed: 2025-12-13.

Bebchuk, Lucian A., Reinier H. Kraakman, and George G. Triantis, "Stock Pyramids, Cross-Ownership, and Dual Class Equity: The Mechanisms and Agency Costs of Separating Control from Cash-Flow Rights," in "Concentrated Corporate Ownership (R.

Morck, ed.),” Harvard Law and Economics Discussion Paper No. 249, January 2000, pp. 295–315.

Buterin, Vitalik, “Ethereum: A Next-Generation Smart Contract and Decentralized Application Platform,” https://ethereum.org/content/whitepaper/whitepaper-pdf/Ethereum_Whitepaper_-_Buterin_2014.pdf 2014. Accessed: 2026-01-06.

Carre, Sylvain and Franck Gabriel, “Liquid Staking: When Does It Help?,” September 2024.

Chemla, Gilles and Christopher A. Hennessy, “Skin in the Game and Moral Hazard,” *The Journal of Finance*, 2014, *69* (4), 1597–1641.

CoinGecko, “StakeWise (SWISE) historical data,” https://www.coingecko.com/en/coins/stakewise 2025. Accessed: 2026-01-02.

CoinGecko (ETH USD price), “ETH/USD historical price,” https://www.coingecko.com/en/coins/ethereum/history 2025. Accessed: 2026-01-01.

Cong, Lin William, Zhiheng He, and Ke Tang, “The Tokenomics of Staking,” April 2025.

DeFiLlama (ether.fi), “Protocol ether.fi (TVL) historical price,” https://defillama.com/protocol/ether.fi 2025. Accessed: 2026-01-01.

DeFiLlama (Lido), “Protocol Lido (TVL) historical price,” https://defillama.com/protocol/lido 2025. Accessed: 2026-01-01.

DeFiLlama (Rocket Pool), “Protocol Rocket Pool (TVL) historical price,” https://defillama.com/protocol/rocket-pool 2025. Accessed: 2026-01-01.

DeFiLlama (StakeWise), “Protocol StakeWise (TVL) historical price,” https://defillama.com/protocol/stakewise 2025. Accessed: 2026-01-01.

Ethereum Foundation, "Attack and defense," https://ethereum.org/developers/docs/consensus-mechanisms/pos/attack-and-defense/ 2026. Accessed: 2026-01-12.

Ethereum Foundation, "Gasper," https://ethereum.org/developers/docs/consensus-mechanisms/pos/gasper/ 2026. Accessed: 2026-01-12.

Ethereum Foundation, "Proof-of-stake (PoS)," https://ethereum.org/developers/docs/consensus-mechanisms/pos/ 2026. Accessed: 2026-01-12.

ether.fi, "Dapp walkthrough," https://etherfi.gitbook.io/etherfi/node-operators/dapp-walkthrough 2026. Accessed: 2026-01-01.

ether.fi, "Node operators guide," https://etherfi.gitbook.io/etherfi/node-operators/node-operators-guide 2026. Accessed: 2026-01-01.

ether.fi, "Technical Documentation," https://etherfi.gitbook.io/etherfi/ether.fi-whitepaper/technical-documentation 2026. Accessed: 2026-03-27.

Gersbach, Hans, Akaki Mamageishvili, and Manvir Schneider, "Staking Pools on Blockchains," October 2022.

Gogol, Krzysztof, Benjamin Kraner, Malte Schlosser, Tao Yan, Claudio Tessone, and Burkhard Stiller, "Empirical and Theoretical Analysis of Liquid Staking Protocols," January 2024.

Gogol, Krzysztof, Yaron Velner, Benjamin Kraner, and Claudio Tessone, "SoK: Liquid Staking Tokens (LSTs) and Emerging Trends in Restaking," December 2024.

Heimbach, Lioba, Eric Schertenleib, and Roger Wattenhofer, "DeFi Lending During The Merge," August 2023.

Irresberger, Felix and Ruomei Yang, "Coin Concentration of Proof-of-Stake Blockchains," *Economics Letters*, August 2023, *229*, 111219.

Jeong, Seungwon (Eugene), "Centralized Decentralization: Does Voting Matter? Simple Economics of the DPoS Blockchain Governance," 2020.

Jermann, Urban J., "Optimal Issuance for Proof-of-Stake Blockchains," September 2024.

John, Kose, Thomas J. Rivera, and Fahad Saleh, "Equilibrium Staking Levels in a Proof-of-Stake Blockchain," November 2021.

Lehar, Alfred, Christine A. Parlour, and Kathy Yuan, "Liquid Staking," March 2025. Preliminary and incomplete; comments welcome.

Lido Finance, "Community Staking Module: a new era for solo stakers," https://blog.lido.fi/community-staking-module-new-era-for-solo-stakers/ 2024. Accessed: 2026-01-01.

Lido Finance, "Community Staking Module," https://lido.fi/csm 2025. Accessed: 2026-01-01.

Lido Finance, "Curated Module: general overview," https://docs.lido.fi/guides/curated-module/general-overview 2025. Accessed: 2025-12-13.

Lido Finance, "Staking Router (contracts)," https://docs.lido.fi/contracts/staking-router/ 2025. Accessed: 2026-01-01.

Lido Finance, "Lido tokens integration guide," https://docs.lido.fi/guides/lido-tokens-integration-guide 2026. Accessed: 2026-01-01.

Lido Finance, "NodeOperatorsRegistry (contracts)," https://docs.lido.fi/contracts/node-operators-registry 2026. Accessed: 2026-01-01.

Pennacchi, George G., "Loan Sales and the Cost of Bank Capital," *The Journal of Finance*, 1988, *43* (2), 375–396.

Rocket Pool, "8-ETH Bonded Minipools," https://docs.rocketpool.net/guides/atlas/lebs 2026. Accessed: 2026-01-02.

Rocket Pool, "Fee Distributors and the Smoothing Pool," https://docs.rocketpool.net/node-staking/fee-distrib-sp 2026. Accessed: 2026-03-27.

Rocket Pool, "A Node Operator's Responsibilities," https://docs.rocketpool.net/guides/node/responsibilities.html#rocket-pool-node-operators 2026. Accessed: 2026-01-02.

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

StakeWise, "osToken: how osToken works," https://docs.stakewise.io/docs/ostoken/how-ostoken-works 2025. Accessed: 2026-01-01.

StakeWise Governance, "SWIP-24: Establish a policy for enabling 100% osETH minting in select Vaults," https://forum.stakewise.io/t/swip-24-establish-a-policy-for-enabling-100-oseth-minting-in-select-vaults/1724 2024. Accessed: 2026-01-27.

Tang, Dunzhe, Ping He, Zhongjie Fan, and Yujie Wang, "Pool Competition and Centralization in PoS Blockchain Network," *Applied Economics*, November 2024, *56* (52), 6564–6583.

Tzinas, Apostolos and Dionysis Zindros, "The Principal–Agent Problem in Liquid Staking," in "International Conference on Financial Cryptography and Data Security" Springer 2023, pp. 456–469.

I Table construction

Table 1 reports selected institutional features, capital structure ratios and design dimensions for four major liquid staking protocols on Ethereum, using data accessed on 2025-12-13. Centralized or custodial LSTs such as WBETH (Binance) and cbETH (Coinbase) are excluded because they are less relevant for the attack profile studied in the paper. Where a protocol has multiple operator modules with distinct capital requirements, the table reports separate rows.

Column definitions. Token records the liquid staking token traded in secondary markets. Type classifies claim origination: compounding denotes modules where users enter at par and claim value evolves through rebasing or exchange-rate appreciation; discounting denotes modules where claims are originated through collateralized minting against a retained collateral slice. TVL is total value locked in millions of ETH. Share is the percentage of total staked ETH (approximately 35.69M from the Beacon Chain). Each capital ratio is per unit of ETH managed for validation. Bond records formal operator bond requirements. Outside-bond capital records required operator capital outside formal bond definitions. Res. records mandatory non-operator reserve or overcollateralization layers. Total is the sum of Bond, outside-bond capital and Res. Perm. indicates whether operator admission is curated or governance-gated. Fee records the headline reward skim or operator commission. 1st-loss records the denomination of the primary mandatory first-loss layer.

Lido. The table covers two modules: the Curated Module and the Community Staking Module (CSM), selected to contrast a permissioned operator set with no mandatory capital layers against a permissionless entry path with explicit bond requirements (Lido Finance, 2024). While Lido's staking architecture can accommodate additional modules via the Staking Router (Lido Finance, 2025c), these two are sufficient for the table's purpose. Both modules are categorized as compounding because users enter at par and stETH value evolves through rebase and share-rate mechanics. The Curated Module is permissioned

and records zeros because the cited documentation does not state a mandatory per-validator capital requirement in the three-layer form used here. The CSM is permissionless; Bond is computed from the first-validator requirement on the CSM page (2.4 ETH) divided by 32 ETH managed, giving 0.075 (Lido Finance, 2025a). The schedule is tiered, so marginal Bond for additional validators is lower. The bond can be posted in ETH, stETH, or wstETH but is held as stETH shares (Lido Finance, 2026a, 2025a). Fee is 10% for both modules. The TVL and Share figures represent the total across both modules.

ether.fi. The cited operator documentation emphasizes permissioning through whitelisting and an auction mechanism for assigning additional validators (ether.fi, 2026a,b). Validators run on Distributed Validator Technology (DVT) clusters via SSV Network: each validator key is split across approximately seven operators using threshold cryptography, with a five-of-seven signing threshold, so that up to two operators can fail without affecting validator duties and no single operator can unilaterally cause a slashing event (ether.fi, 2026b). The module is categorized as compounding because users enter at par and eETH accrues value through rebase and share-rate updating. The table records zeros across the capital layers because the cited documentation does not specify a mandatory per-validator capital requirement in this form. Fee is protocol-set and governance-adjustable (ether.fi, 2026c).

Rocket Pool. The cited documentation emphasizes permissionless entry paired with explicit capital requirements: node operators contribute bonded ETH alongside pooled ETH from the deposit pool to create validators and also post RPL collateral (Rocket Pool, 2026c,a). The module is categorized as compounding because users enter through pooled staking and rETH value accrues through an increasing exchange rate. Bond is 8/32 = 0.25 under LEB8. Outside-bond capital adds minimum required RPL collateral equivalent to 2.4 ETH per validator, so 2.4/32 = 0.075; Total is therefore $(8 + 2.4)/32 \approx 0.325$ (Rocket Pool, 2026a,c). Fee is a 5–14% operator commission on pooled ETH (Rocket Pool, 2026b).

StakeWise. The table distinguishes two vault tiers that differ in which layer supplies first-loss capital. Both tiers are categorized as discounting because osETH is originated through collateralized minting against a retained collateral slice rather than through pooled par issuance. In standard vaults (90% LTV), outside-bond capital includes a 1-ETH onboarding collateral commitment, giving 1/32 = 0.031. Res. is the mandatory 10% buffer implied by the 90% LTV cap, giving 3.2/32 = 0.10 (StakeWise, 2025). Total is therefore $(1+3.2)/32 \approx 0.131$. In DAO-approved 100% LTV vaults, the 10% reserve layer is removed and the operator must lock 5,000,000 SWISE as slashing insurance (StakeWise Governance, 2024). The SWISE component equals about 0.000860 using SWISE/ETH on 2025-12-13 and the 10,000 ETH minimum TVL threshold (StakeWise, 2025; CoinGecko, 2025). Adding the 1-ETH onboarding collateral gives outside-bond capital of about 0.032 and Total about 0.032 in the 100% LTV row. The 100% LTV tier is permissioned: only vaults meeting DAO-defined criteria (minimum 10,000 ETH TVL, above-median performance, maximum 5% fee, latest vault version and the SWISE bond) are approved (StakeWise Governance, 2024). Fee is vault-set in the standard tier and capped at 5% in the 100% LTV tier. The TVL and Share figures represent the total across both tiers.

Data sources. The share denominator uses total staked ETH (approximately 35.69M, accessed on 2025-12-13) from the Beacon Chain (Beaconcha, 2025). TVL and Share are computed from DeFiLlama protocol TVL (USD) using the DeFiLlama historical ETH/USD price at the same timestamp (DeFiLlama , Lido,e,R,S; CoinGecko , ETH USD price).

II Proofs

Proof of Proposition 1

If an attack occurs, slashing destroys γS units of ETH on the stake controlled by the attacker. By assumption, the protocol can reimburse the pool using at most the collateral amount xS, so the residual loss passed through to LST holders is $U = \max\{0, \gamma S - xS\}$, which yields

equation (1). With backing M^L and a pre-event token supply of $L = (1-x)M^L$, each unit of the LST is a pro-rata claim on the non-collateral slice $(1-x)M^L$, so after the loss the backing per token is

$$\min\left\{\frac{(1-x)M^L - U}{L},\ 1\right\}.$$

Using $L = (1-x)M^L$, we have

$$\frac{(1-x)M^L - U}{L} = 1 - \frac{U}{(1-x)M^L}.$$

Substituting $U = \max\{0, (\gamma - x)S\}$ yields

$$1 - \frac{U}{(1-x)M^L} = 1 - \max\{0,\ \gamma - x\}\left(\frac{1}{1-x}\right)\frac{S}{M^L},$$

which gives equation (6).

Proof of Lemma 1

Case (i): no-attack candidate ($q = 0$). In a pure-strategy equilibrium with $q = 0$, the $t = 2$ backing is $V(0) = 1$, so competitive pricing implies $p = 1$. This outcome is an equilibrium if the attacker does not deviate to attacking at $p = 1$, i.e. if $\Delta\Pi_{\text{liquid}}(1) \le 0$. By the definition of $p_{\text{profitable}}$, this is equivalent to $1 \le p_{\text{profitable}}$.

Case (ii): attack candidate ($q = 1$). In a pure-strategy equilibrium with $q = 1$, the $t = 2$ backing is $V(1) = b$, so competitive pricing implies $p = b$. This outcome is an equilibrium if the attacker optimally attacks at that price, i.e. if $\Delta\Pi_{\text{liquid}}(b) \ge 0$, which is equivalent to $b \ge p_{\text{profitable}}$.

Case (iii): intermediate region and mixed equilibrium. When $b < p_{\text{profitable}} < 1$, neither pure-strategy profile is an equilibrium: if $A = 1$, then $p = b < p_{\text{profitable}}$ and the attack is not profitable, while if $A = 0$, then $p = 1 > p_{\text{profitable}}$ and the attack is profitable. For a

mixed-strategy equilibrium, let q be the attack probability. Competitive pricing implies

$$p = (1 - q) \cdot 1 + q \cdot b = 1 - q(1 - b).$$

The attacker must be indifferent between attacking and not attacking, so the indifference condition (7) must hold, and substituting yields $q = \frac{1-p_{\text{profitable}}}{1-b}$. This q lies in $(0, 1)$ exactly when $b < p_{\text{profitable}} < 1$.

Proof of Proposition 2

The objects $p_{\text{profitable}}$ and b depend on x through equation (5) and equation (6). We write $p_{\text{profitable}}(x)$ and $b(x)$ to emphasize this dependence.

By Lemma 1, case (i) holds exactly when $p_{\text{profitable}}(x) \geq 1$. When $x < \gamma$, we have $\max\{0,\, x - \gamma\} = 0$, so equation (5) becomes $p_{\text{profitable}}(x) = \left(\frac{1}{1-x}\right)\left[1 + (1 - \tau)\lambda - \frac{a}{S}\right]$, which implies $p_{\text{profitable}}(x) \geq 1$ if and only if $x \geq x_0$. When $x \geq \gamma$, we have $\max\{0,\, x - \gamma\} = x - \gamma$, so equation (5) becomes $p_{\text{profitable}}(x) = \left(\frac{1}{1-x}\right)\left[1 + (1 - \tau)\lambda - (1 - d)(x - \gamma) - \frac{a}{S}\right]$, which implies $p_{\text{profitable}}(x) \geq 1$ if and only if $(1 - \tau)\lambda + dx + (1 - d)\gamma \geq \frac{a}{S}$, yielding the second condition in (i).

Lemma 1(ii) holds when $b(x) \geq p_{\text{profitable}}(x)$. When $x \geq \gamma$, we have $b(x) = 1$ so $b(x) \geq p_{\text{profitable}}(x)$ is equivalent to $1 \geq p_{\text{profitable}}(x)$, i.e., to $(1 - \tau)\lambda + dx + (1 - d)\gamma < \frac{a}{S}$, yielding the first condition in (ii). When $x < \gamma$, substituting the corresponding expressions for $b(x)$ and $p_{\text{profitable}}(x)$ from the discount model yields a linear inequality in x whose unique cutoff is x_1. Intersecting $b(x) \geq p_{\text{profitable}}(x)$ with $p_{\text{profitable}}(x) < 1$ (equivalently $x < x_0$) yields $x < \min\{x_0, x_1\}$, giving the second condition in (ii). When $x < \gamma$, we have $p_{\text{profitable}}(x) = \left(\frac{1}{1-x}\right)\left[1 + (1 - \tau)\lambda - \frac{a}{S}\right]$ and $b(x) = 1 - \left(\frac{S}{M^L}\right)\left(\frac{\gamma - x}{1-x}\right)$. Substituting these expressions into $b(x) \geq p_{\text{profitable}}(x)$ and rearranging yields the cutoff $x \leq x_1$.

Lemma 1(iii) requires $b(x) < p_{\text{profitable}}(x) < 1$. This cannot occur when $x \geq \gamma$ because then $b(x) = 1$. In the bond-limited region $x < \gamma$, the conditions $p_{\text{profitable}}(x) < 1$ and $b(x) < p_{\text{profitable}}(x)$ are equivalent to $x < x_0$ and $x > \min\{x_0, x_1\}$, respectively, which

gives $\min\{x_0, x_1\} < x < \max\{x_0, x_1\}$. The mixed-equilibrium formula for $q(x)$ is the one in Lemma 1(iii). When $x < \gamma$, substituting $p_{\text{profitable}}(x) = \left(\frac{1}{1-x}\right)\left[1 + (1-\tau)\lambda - \frac{a}{S}\right]$ and $b(x) = 1 - \left(\frac{S}{M^L}\right)\left(\frac{\gamma - x}{1-x}\right)$ into $q = \frac{1 - p_{\text{profitable}}}{1-b}$ yields $q = \frac{\frac{a}{S} - (1-\tau)\lambda - x}{\frac{S}{M^L}(\gamma - x)}$.

Proof of Corollary 1

Fix $x \leq 1$. If $x < \gamma$, then $\min(\gamma, x) = x$ and

$$a^*_{\text{liquid}}(x) = S\Big((1-\tau)\lambda + x\Big).$$

Since $x < \gamma$ and $x \leq 1$, we have $x \leq \min\{\gamma, 1\}$. Moreover, for any $d \in [0,1]$ we have $\min\{\gamma, 1\} \leq d + (1-d)\gamma$. Therefore,

$$a^*_{\text{liquid}}(x) = S\Big((1-\tau)\lambda + x\Big) \leq S\Big((1-\tau)\lambda + \min\{\gamma, 1\}\Big) \leq S\Big(\lambda + d + (1-d)\gamma\Big) = a^*_{\text{traditional}}.$$

If $x \geq \gamma$, then $\min(\gamma, x) = \gamma$ and

$$a^*_{\text{liquid}}(x) = S\Big((1-\tau)\lambda + dx + (1-d)\gamma\Big) \leq S\Big(\lambda + d + (1-d)\gamma\Big) = a^*_{\text{traditional}},$$

where the inequality uses $(1-\tau)\lambda \leq \lambda$ and $dx \leq d$.

Proof of Proposition 3

Fix (q, M^O) with $q \in [0,1]$ and $M^O > 0$. Equations (15) and (16) pin down M^S, S, and M^L. Equations (6), (10) and (5) then determine b, p and $p_{\text{profitable}}$.

In Case (i), when $q = 0$, equation (10) implies $p = \frac{1}{1+\mu}$. A no-attack operating equilibrium therefore exists if and only if some $M^O > 0$ satisfies equation (14) at $q = 0$ and the attacker does not deviate, which is equivalent to $p_{\text{profitable}} \geq \frac{1}{1+\mu}$.

In Case (ii), when $q = 1$, equation (10) implies $p = \frac{1}{1+\mu} b$. An attack operating equilibrium therefore exists if and only if some $M^O > 0$ satisfies equation (14) at $q = 1$ and

the attacker is willing to attack at that price, which is equivalent to $b \geq (1+\mu)p_{\text{profitable}}$.

In Case (iii), with $q \in (0,1)$, the attacker is indifferent, so the indifference condition (7) must hold. Together with equation (14), this is necessary and sufficient for a mixed operating equilibrium. When $x < \gamma$, we have backing below par $b < 1$ after attack, so combining equations (7) and (10) pins down

$$q = \frac{1-(1+\mu)p_{\text{profitable}}}{1-b}.$$

Hence a mixed equilibrium exists if and only if this value lies in $(0,1)$ and some $M^O > 0$ also satisfies equation (14). When $x \geq \gamma$, equation (6) gives $b = 1$, so equation (10) implies $p = \frac{1}{1+\mu}$ for any $q \in [0,1]$. The mixed condition then reduces to $p_{\text{profitable}} = \frac{1}{1+\mu}$ together with equation (14) and $q \in (0,1)$.

In Case (iv), if no pair (q, M^O) with $M^O > 0$ satisfies Cases (i) to (iii), the residual equilibrium is shutdown with $M^O = 0$.

Proof of Corollary 2

Fix $\mu > 0$ and $x \leq 1$. Let $M_0^O(x) > 0$ denote an operating no-attack candidate scale satisfying honest-operator participation at $q = 0$ with no-attack price $p = \frac{1}{1+\mu}$. Define

$$S_0(x) := \left(\frac{f}{1-f}\right)\left[\alpha + M_0^O(x)\right], \qquad \lambda_0(x) := \frac{k\sqrt{1-f}}{\sqrt{\alpha + M_0^O(x)}}.$$

Under $q = 0$, deterrence requires

$$\left(\frac{1}{1-x}\right)\left[1 + (1-\tau)\lambda_0(x) - (1-d)\max\{0,\, x-\gamma\} - \frac{a}{S_0(x)}\right] \geq \frac{1}{1+\mu}.$$

Characterize the two cutoff objects by

$$a^*_{\text{traditional}} = S_0(x)\Big(\lambda_0(x) + d + (1-d)\gamma\Big),$$

and

$$a^*_{\text{liquid}}(x) = S_0(x)\left(1 + (1-\tau)\lambda_0(x) - (1-d)\max\{0,\, x-\gamma\} - \frac{1-x}{1+\mu}\right), \tag{19}$$

where equation (19) is equivalent to the deterrence inequality above. By Lemma 5, S_0 and λ_0 are independent of x at $q = 0$. The same lemma gives $(1-\tau)\lambda_0 = \mu$, so

$$a^*_{\text{traditional}} - a^*_{\text{liquid}}(x) = S_0(x)\left(\frac{\tau}{1-\tau}\mu + \frac{1-x}{1+\mu} - (1-d)\big(1-\gamma-\max\{0,\, x-\gamma\}\big)\right).$$

If $x < \gamma$, this becomes

$$a^*_{\text{traditional}} - a^*_{\text{liquid}}(x) = S_0(x)\left(\frac{\tau}{1-\tau}\mu + \frac{1-x}{1+\mu} - (1-d)(1-\gamma)\right).$$

This expression is decreasing in x, so on $[0, \gamma)$ it is minimized at $x = \gamma$. If $x \geq \gamma$, then

$$a^*_{\text{traditional}} - a^*_{\text{liquid}}(x) = S_0(x)\left(\frac{\tau}{1-\tau}\mu + (1-x)\left(\frac{1}{1+\mu} - (1-d)\right)\right).$$

If $\frac{1}{1+\mu} \geq 1-d$, this expression is minimized at $x = 1$, where it equals $S_0(x)\frac{\tau}{1-\tau}\mu > 0$. If $\frac{1}{1+\mu} < 1-d$, it is minimized at $x = \gamma$. Therefore a sufficient condition for $a^*_{\text{traditional}} - a^*_{\text{liquid}}(x) \geq 0$ for all $x \in [0, 1]$ is

$$\frac{\tau}{1-\tau}\mu + \frac{1-\gamma}{1+\mu} \geq (1-\gamma)(1-d),$$

which proves the claim.

Proof of Lemma 6

Under equation (10), we have

$$(1+\mu)p = (1-q)\cdot 1 + q\cdot b, \qquad (1+\mu)(1-p) = \mu + q\big(1-b\big),$$

Substituting this identity into the rearranged participation condition

$$(1-qd)(1-\tau)\lambda - (1-p)(1+\mu)(1-x) - dqx = \mu x$$

gives

$$(1-qd)(1-\tau)\lambda = \mu + dqx + q(1-x)\big(1-b\big) \geq \mu.$$

If $d > 0$ then $1 - qd < 1$, and if $x < \gamma$ then $b < 1$ so the second term on the right-hand side is strictly positive. In either case,

$$(1-\tau)\lambda > \mu.$$

Since λ is strictly decreasing in M^O, this implies $M^O < M_0^O$. In the remaining knife-edge case ($d = 0$ and $x \geq \gamma$), we have $b = 1$ and the condition reduces to $(1-\tau)\lambda = \mu$, so $M^O = M_0^O$.

Proof of Proposition 4

Starting from the rearranged participation condition

$$(1-qd)(1-\tau)\lambda - (1-p)(1+\mu)(1-x) - dqx = \mu x,$$

use

$$(1+\mu)(1-p) = \mu + q(1-b),$$

which follows from equation (10). Then

$$(1-qd)(1-\tau)\lambda = \mu x + dqx + (1-x)\big[\mu + q(1-b)\big] = \mu + qdx + q(1-x)\big(1-b\big).$$

Dividing by $1 - qd$ and moving all terms to the left-hand side gives

$$(1-\tau)\lambda - \frac{\mu}{1-qd} - \frac{qdx}{1-qd} - \frac{q(1-x)\big(1-b\big)}{1-qd} = 0,$$

This is the honest-operator participation condition (14), written as a surplus decomposition. The term qdx lowers surplus because a larger fraction of operator value remains exposed to depreciation. The term $q(1-x)\big(1-b\big)$ is the expected loss from uncovered slashing borne through the liquid portion of the stake, and it falls as x rises because the liquid share shrinks and, when $x < \gamma$, backing improves. Since λ is decreasing in M^O, an upward shift in surplus supports a higher equilibrium M^O.

The discount-deterrence channel does not appear as a separate additive term because it operates through equilibrium q. When $x < \gamma$, a higher bond ratio raises b and weakens the price response to attack beliefs. This can sustain a higher equilibrium attack probability, which lowers participation surplus through the terms already displayed above. If $x \geq \gamma$, then uncovered slashing is fully covered, so $b = 1$, the socialized-slashing term vanishes, and the discount-deterrence channel is absent. The overall effect of x on equilibrium honest stake is therefore not signed in general.

Proof of Lemma 7

Using

$$\lambda = \frac{k\sqrt{1-f}}{\sqrt{\alpha + M^O}}, \qquad M^L = \left(\frac{1}{1-f}\right)\left[f\alpha + M^O\right],$$

the protocol payoff becomes

$$\Pi^P = \frac{\tau k\sqrt{1-f}}{1-f} \cdot \frac{f\alpha + M^O}{\sqrt{\alpha + M^O}}.$$

Differentiate this expression with respect to M^O:

$$\frac{d\Pi^P}{dM^O} = \frac{\tau k\sqrt{1-f}}{1-f} \cdot \frac{\alpha(2-f) + M^O}{2(\alpha + M^O)^{3/2}}.$$

Since $f \in (0,1)$, $\alpha \geq 0$, and $M^O \geq 0$, the numerator $\alpha(2-f) + M^O$ is strictly positive, so $\frac{d\Pi^P}{dM^O} > 0$.

Proof of Proposition 6

Using equation (15), we have

$$M^L = \left(\frac{1}{1-f}\right)\left[f\alpha + M^O\right], \qquad S = \left(\frac{f}{1-f}\right)\left[\alpha + M^O\right].$$

Therefore $M^L - S = M^O$, so $M^L \geq S$ holds whenever $M^O \geq 0$.

Next, if $x \leq \gamma$ then equation (5) implies

$$p_{\text{profitable}} = \frac{1}{1-x}\left[1 + (1-\tau)\lambda - \frac{a}{S_0}\right].$$

Lemma 5 gives $(1-\tau)\lambda = \mu$ at the no-attack operating candidate $M^O = M_0^O$, so

$$p_{\text{profitable}} = \frac{1}{1-x}\left[1 + \mu - \frac{a}{S_0}\right].$$

The deterrence condition $p_{\text{profitable}} \geq \frac{1}{1+\mu}$ is equivalent to

$$\frac{1}{1-x}\left[1 + \mu - \frac{a}{S_0}\right] \geq \frac{1}{1+\mu},$$

which rearranges to $x \geq 1 - (1+\mu)^2 + \frac{(1+\mu)a}{S_0}$.

If $x > \gamma$, then equation (5) gives

$$p_{\text{profitable}} = \frac{1}{1-x}\left[1 + (1-\tau)\lambda - (1-d)(x-\gamma) - \frac{a}{S_0}\right].$$

Using Lemma 5 again, substitute $(1-\tau)\lambda = \mu$ and obtain

$$\frac{1}{1-x}\left[1 + \mu - (1-d)(x-\gamma) - \frac{a}{S_0}\right] \geq \frac{1}{1+\mu},$$

which is the stated condition.

Online Appendix: For Online Publication Only

A USD-denominated attacker problem

This section studies a USD-denominated variant of the attacker's problem in which an attack induces a proportional depreciation of the USD/ETH exchange rate. The purpose is to show that, once the continuation value is solved in closed form, the profitability cutoff has the same structure as in the ETH-denominated formulation, with the remaining difference coming from a constant scaling term induced by discounting.

A USD-denominated formulation must specify what happens after a no-attack history, because the possibility of attack can affect the USD/ETH exchange rate before an attack is realized. In a finite-horizon formulation, a no-attack history eventually reaches a terminal state. If that state is interpreted as the honest outcome with no remaining attack risk, USD/ETH exchange rate returns to its baseline level, and that terminal normalization propagates backward, mechanically generating nonstationary pre-terminal price dynamics. To avoid this artifact, we use an infinite-horizon repeat-until-attack structure. Under a fixed belief q, the same USD/ETH pricing rule then applies in every no-news period, so the attacker's continuation value is stationary.

We therefore consider an infinite-horizon economy that repeats each period until the first attack. Fix a market belief q that an attack occurs in a given period, and let $\beta \in (0, 1)$ denote the discount factor. In each period, the LST trades competitively at price p (ETH per LST), while ETH trades at $P_1(q)$ USD. If an attack occurs, the ETH price drops to $P_2(q) = (1 - d)P_1(q)$ and the game ends. If no attack occurs, the ETH price remains $P_1(q)$ and the economy continues with the same belief q. Because the model has no private information, observing no attack does not change beliefs or prices.

Let $V(q)$ denote the attacker's continuation value in USD at the beginning of a period under belief q. In each period, the attacker chooses an action $A \in \{0, 1\}$, where $A = 1$ denotes attack. Conditional on q, the stationary Bellman equation is

$$V(q) = \max \left\{ V^1(q),\, V^0(q) \right\},$$

where $V^1(q)$ is the value from attacking in the current period, and $V^0(q)$ is the value from not attacking so that the game continues with the same belief q.

If there's an attack, the attacker liquidates its $(1 - x)S$ LST claims at price p, obtains the private benefit a (in ETH terms), and retains the remaining xS as staked collateral. Slashing on the attacker-controlled stake destroys γS units of ETH. Because collateral absorbs losses up to xS, the remaining collateral value after slashing is $\max\{0,\, x - \gamma\}S$, which is exposed

to the post-attack ETH price $P_2(q) = (1-d)P_1(q)$. The private benefit a is not exposed to this depreciation channel, so its USD value is $P_1(q)a$.[1] The attack value therefore has no continuation term:

$$V^1(q) = -P_1(q)S + P_1(q)p(1-x)S + P_1(q)a + P_2(q)\max\{0,\, x-\gamma\}S,$$
$$P_2(q) = (1-d)P_1(q).$$

If the attacker does not attack, it keeps the position, earns the within-period flow payoff, and carries the position into the next period. The no-attack value is therefore

$$V^0(q) = P_1(q)(1-\tau)\lambda S + \beta V(q).$$

In any period, attack is optimal if and only if $V^1(q) \geq V^0(q)$. For a given continuation value $V(q)$, this condition can be written as a cutoff for the secondary-market price p:

$$p \geq p_{\text{profitable}}(q),$$
$$p_{\text{profitable}}(q) \coloneqq \left(\frac{1}{1-x}\right)\left[1 + (1-\tau)\lambda + \frac{\beta V(q)}{P_1(q)S} - \frac{a}{S} - (1-d)\max\{0,\, x-\gamma\}\right]. \tag{A.1}$$

At the indifference point $V^1(q) = V^0(q)$, the Bellman equation can be solved in closed form. Then $V(q) = V^0(q)$, so

$$V(q) = \frac{P_1(q)(1-\tau)\lambda S}{1-\beta}. \tag{A.2}$$

Substituting equation (A.2) into equation (A.1) shows that the cutoff depends on primitives but not on q:

$$p \geq p_{\text{profitable}}, \qquad p_{\text{profitable}} = \left(\frac{1}{1-x}\right)\left[1 + \frac{(1-\tau)\lambda}{1-\beta} - (1-d)\max\{0,\, x-\gamma\} - \frac{a}{S}\right]. \tag{A.3}$$

The comparison with the ETH-denominated formulation is now straightforward. Because $P_1(q)$ scales all within-period dollar payoffs and also scales the continuation value in (A.2), it cancels from the profitability cutoff after substitution. One can therefore normalize

[1] This corresponds to modeling the attack benefit in the ETH-denominated formulation as a rather than $a(1-d)$. If we instead used $a(1-d)$ in the ETH-denominated formulation, the attack value would be $P_1(q)a(1-d)$. There is little difference in the implications.

$P_1(q) = 1$ (so that $P_2(q) = 1 - d$) and interpret the resulting payoffs in ETH terms.

The remaining difference relative to the main text is discounting rather than denomination. In the main text, the attacker's objective is written directly in ETH and d scales down the value of ETH that remains exposed through the event. Here the objective is written in USD and d is interpreted as proportional depreciation of USD/ETH conditional on attack, so rational expectations can in principle move the exchange rate before the attack is realized. But once the continuation value is substituted, the cutoff in equation (A.3) differs from the main-text cutoff only through the opportunity-cost term $\frac{(1-\tau)\lambda}{1-\beta}$, which is the present value of foregone net staking rewards in the stationary infinite-horizon environment. In the main text, the corresponding two-date accounting contributes only the one-period term $(1 - \tau)\lambda$. Up to this discounting adjustment, the dependence of the profitability cutoff on (x, γ, d, a, S) is the same as in equation (5), including the factor $\frac{1}{1-x}$ reflecting that only $(1 - x)S$ liquid claims can be sold.

B Attacker participation and the outside option of traditional staking

The design of the problem in the main text treats the attacker's liquid-staking participation as given and derives the equilibrium attack probability q conditional on entry. A further margin is whether the attacker enters that strategy at all rather than honestly stake the same capital through traditional staking.[2] This additional participation constraint can eliminate some candidate attack branches, but, as the next proposition shows, it does not change the policy ranking when an optimal policy in the baseline problem already implements a no-attack equilibrium. We prove this result in the context of the endogenous-scale model presented in Section 4. An analogous result for the exogenous-scale model in Section 3 follows by the same logic.

To model this margin in the simplest way, suppose the attacker can instead stake the same capital S traditionally and earn the endogenous net staking return λ per unit of ETH. The outside option therefore yields λS.

If the attacker enters the liquid-staking strategy and stays honest within the protocol, his

[2]We do not consider the alternative outside option of attacking through traditional staking, because in the presence of the liquid-staking protocol this strategy is strictly dominated.

net payoff is

$$\tilde{\Pi}^0_{\text{liquid}} = (1-\tau)\lambda\, S.$$

If he attacks, his net payoff is

$$\tilde{\Pi}^1_{\text{liquid}}(p) \coloneqq \tilde{\Pi}^0_{\text{liquid}} + \Delta\Pi_{\text{liquid}}(p),$$

where $\Delta\Pi_{\text{liquid}}(p)$ is given by equation (4). For a candidate equilibrium (q, M^O), equations (15) and (16) determine S and M^L. Equations (1), (6) and (10) then determine U, b and p.

The attacker's ex ante expected payoff from participating in the liquid-staking strategy is therefore

$$\mathbb{E}\big[\tilde{\Pi}_{\text{liquid}} \mid q, M^O\big] = \tilde{\Pi}^0_{\text{liquid}} + q\,\Delta\Pi_{\text{liquid}}(p).$$

Entry in the liquid-staking protocol requires

$$\mathbb{E}\big[\tilde{\Pi}_{\text{liquid}} \mid q, M^O\big] \;\geq\; \lambda\, S.$$

Proposition B.1 (No-attack optimality under attacker non-participation)**.** *If a policy that is optimal in the baseline design problem implements a no-attack operating equilibrium, then it remains optimal when the attacker can instead opt out and stake honestly in traditional staking.*

Proof. Fix a policy that is optimal in the baseline problem and implements a no-attack operating equilibrium. Adding the participation margin may remove some positive-attack candidates because the attacker may prefer the traditional-staking outside option. If that happens under this policy, the protocol remains at the same no-attack allocation and earns the same payoff. The extra constraint therefore prunes the equilibrium set without creating any policy that yields higher payoff than the optimal no-attack policy in the baseline problem. Hence that policy remains optimal. ■

C Compounding model

This section develops the compounding version of the main model. In this version, the liquid claim is originated directly at the protocol level. Depositors send ETH to the liquid-staking pool and receive newly issued LST at par, so they are the ultimate holders of the liquid claim and play the same conceptual role as secondary-market LST buyers in the discount model.

Honest operators and the attacker, by contrast, obtain control over stake by posting the required bond as collateral, rather than by first supplying the full stake and then distributing or selling claims against it. These institutional differences change the accounting, but not the underlying economic objects: uncovered loss, attacker incentives, and depositor participation still rest on equations (1), (4), and (10). Claims are issued at par, rewards are assigned through rebase, and attack-state continuation values are therefore pinned down by how pooled ETH and LST supply evolve between dates 1 and 2. The discussion follows that logic. We first map depositor participation scale into staking, attack, and bonding quantities; next we define the state-transition system that generates backing; and finally we characterize equilibrium, deterrence, and protocol design. Let M_t^L denote pooled ETH at date t and L_t the LST supply, so backing per token is M_t^L/L_t. At origination the protocol issues claims at par:

$$L_1 = M_1^L.$$

Let M^O denote honest operators' bonded capital and M^A the attacker's bonded capital. The bond-ratio condition is

$$\frac{M^O + M^A}{M^L} = x. \tag{C.1}$$

The attacker participates as an operator, posting bond capital and earning operator fee rewards through the same rebase mechanism as honest operators. Since each unit of bonded capital supports $1/x$ units of operated stake, the attacker controls staked ETH equal to $\frac{1}{x}M^A$. To reach the attack threshold at minimum capital outlay, equilibrium requires

$$\frac{1}{x}M^A = S \quad \Longleftrightarrow \quad M^A = xS. \tag{C.2}$$

Let τ_N denote the operator fee share, τ_P the protocol-treasury fee share, and $\tau := \tau_N + \tau_P$ the aggregate fee share.

The uncovered-loss expression (1) is unchanged and is recorded here for later reference:

$$U = \gamma S - \min(\gamma S,\ xS) = \max\{0,\ (\gamma - x)S\}\,.$$

C.1 Attack-scale accounting

This subsection relates depositor ETH M^D to the aggregate quantities that later determine pricing, deterrence and feasibility. Total ETH pooled in the liquid-staking protocol is

$$M^L := M^D + M^O + M^A.$$

To derive total staking and the attack scale, let M^T denote traditionally staked ETH and let M be total ETH supply. Write the traditional-staking share as $M^T/M = \alpha$ and normalize $M = 1$. Total staked ETH is then

$$M^S = M^T + M^L = \alpha + M^D + M^O + M^A.$$

Assume the attack scale is a fixed fraction of total staked ETH:

$$S = fM^S, \qquad f \in (0, 1).$$

Substituting equations (C.1) and (C.2) yields

$$M^L = \left(\frac{1}{1-x}\right) M^D, \qquad M^S = \alpha + \left(\frac{1}{1-x}\right) M^D, \qquad S = f\alpha + \left(\frac{f}{1-x}\right) M^D. \tag{C.3}$$

Thus, once M^D is fixed, aggregate scale and attack scale are fixed as well. The corresponding attacker and honest-operator bond positions are

$$M^A = xS = xf\alpha + \left(\frac{x}{1-x}\right) fM^D,$$

$$M^O = \left(\frac{x}{1-x}\right)(1-f)M^D - xf\alpha.$$

The endogenous staking return is

$$\lambda = \frac{k}{\sqrt{M^S}}, \tag{C.4}$$

where $k > 0$ is the protocol issuance constant (Jermann, 2023, 2024).

The remaining accounting restriction is that honest operators must be able to supply a nonnegative bond position. Using the bond-ratio condition together with the expressions

above,

$$M^O = xM^L - M^A = \frac{x}{1-x}(1-f)M^D - xf\alpha,$$

where the second equality uses $M^L = \frac{1}{1-x}M^D$ and $S = f\alpha + \frac{f}{1-x}M^D$. Imposing feasibility, $M^O \geq 0$, gives

$$M^D \geq \frac{f(1-x)}{1-f}\alpha. \tag{C.5}$$

C.2 LST state transitions and backing

Let L_t^D, L_t^O, L_t^A, and L_t^P denote the date-t LST holdings of depositors, honest operators, the attacker, and the protocol treasury. Market clearing requires

$$L_t = L_t^D + L_t^O + L_t^A + L_t^P.$$

The core accounting difference relative to the discount model is that continuation payoffs are not summarized by a date-1 price. They are generated by a sequence of date-2 transformations of the pool and the claim supply. This subsection therefore writes the mechanism in the same order in which it operates: first the pool is updated, then attack losses are allocated through burn, then fee claims are rebased into new LST, and finally backing is computed from the resulting pool-to-supply ratio. Only active operators generate operator rewards, so in attack states the attacker is excluded from that reward allocation.

In step (i), the pool updates according to

$$M_2^L = M_1^L + \Delta M_2^L,$$

with

$$\Delta M_2^L := \left(\frac{L_1^O + \mathbf{1}_{\{A_2=0\}}L_1^A}{L_1^O + L_1^A}\right)\lambda M_1^L - \mathbf{1}_{\{A_2=1\}}\gamma S_1,$$

where $\mathbf{1}_{\{\cdot\}}$ stands for the indicator function. In step (ii), penalties and burn satisfy

$$B_2^{\text{penalty}} = \mathbf{1}_{\{A_2=1\}}\gamma S_1 \frac{L_1}{M_2^L}, \qquad B_2^{\text{burn}} = \min\{B_2^{\text{penalty}}, L_1^A\},$$

$$\tilde{L}_2 = L_1 - B_2^{\text{burn}}, \qquad \tilde{L}_2^A = L_1^A - B_2^{\text{burn}}, \qquad \tilde{L}_2^j = L_1^j \text{ for } j \in \{D, O, P\}.$$

In step (iii), positive net pool growth is split into operator and treasury fee claims and

assigned through rebase:

$$I_2 = \mathbf{1}_{\{\Delta M_2^L > 0\}} \frac{(\tau_N + \tau_P)\Delta M_2^L}{M_2^L - (\tau_N + \tau_P)\Delta M_2^L} \tilde{L}_2,$$

$$I_2^j = \mathbf{1}_{\{\Delta M_2^L > 0\}} \frac{\tau_j \Delta M_2^L}{M_2^L - (\tau_N + \tau_P)\Delta M_2^L} \tilde{L}_2, \qquad j \in \{O, A, P\},$$

$$L_2 = \tilde{L}_2 + I_2, \qquad L_2^j = \tilde{L}_2^j + I_2^j \text{ for } j \in \{O, A, P\}, \qquad L_2^D = \tilde{L}_2^D.$$

Here I_2 is the total quantity of newly issued LSTs, while I_2^j is the portion assigned to group j. The denominator converts the ETH fee share into LST units at the post-update backing rate. In step (iv), date-2 backing is

$$b = \frac{M_2^L}{L_2}.$$

Within the operator share, τ_N is split across honest operators and the attacker in proportion to their post-burn operator holdings, except that in attack states ($A_2 = 1$) the attacker receives no operator reward. Therefore,

$$\tau_O = \left(\frac{\tilde{L}_2^O}{\tilde{L}_2^O + \mathbf{1}_{\{A_2=0\}} \tilde{L}_2^A} \right) \tau_N, \qquad \tau_A = \left(\frac{\mathbf{1}_{\{A_2=0\}} \tilde{L}_2^A}{\tilde{L}_2^O + \mathbf{1}_{\{A_2=0\}} \tilde{L}_2^A} \right) \tau_N,$$

so that $\tau_N = \tau_O + \tau_A$.

The burn rule in step (ii) is intentionally conservative. It converts the slashing loss γS_1 from ETH into burned LST using the factor L_1/M_2^L, namely the ratio of pre-burn supply to the post-slashing pool. A more neutral design would instead use the post-burn supply L_2, defining $B_2^{\text{penalty}} = \gamma S_1 \cdot L_2/M_2^L$. Since $L_2 = L_1 - B_2^{\text{burn}}$ and $B_2^{\text{burn}} = \min\{B_2^{\text{penalty}}, L_1^A\}$, that alternative requires solving a fixed point in B_2^{penalty}.

When $B_2^{\text{penalty}} \le L_1^A$, the fixed-point solution is $B_2^{\text{penalty}} = \gamma S_1 L_1/(M_2^L + \gamma S_1)$, which is strictly less than $\gamma S_1 L_1/M_2^L$. Under that neutral rule, the ETH value of the burned LSTs at the post-burn backing rate equals exactly γS_1, so the burn just offsets the slashing loss. Under the adopted L_1 formula, more LST is burned, so the attacker is over-penalized and the excess is transferred to remaining holders through higher post-burn backing. Because this lowers the attacker's attack-state continuation payoff $\left[bL_2^A\right]_{A_2=1}$, it raises the deterrence threshold a^*_{liquid} relative to the neutral alternative. The control-exposure wedge is therefore, if anything, larger than the one reported in the baseline calculations, so the paper's conclusion that liquid staking weakens deterrence remains a conservative statement.

C.3 Attacker's problem

The attacker's decision problem has the same attack-versus-no-attack logic as in the main text, but the relevant continuation objects now come from the transition system in Section C.2. In particular, all continuation effects are summarized by date-2 backing b and the attacker's date-2 claim holdings L_2^A.

The attacker posts bond capital M_1^A at date 1 and receives LST holdings that may subsequently be burned or rebased. Let L_2^A denote the attacker's date-2 LST holdings and let b denote date-2 backing per LST. Conditional on attack, the purchasing power of one unit of ETH falls by $d \in [0, 1]$. Taking these state-contingent continuation values as given, the attacker chooses between attack and no attack to maximize expected payoff

$$\Pi_{\mathrm{A}}(M_1^A) := \mathbb{E}_1\left[\mathbf{1}_{\{A_2=1\}}a + (1 - d\mathbf{1}_{\{A_2=1\}})bL_2^A - M_1^A\right].$$

Using $M_1^A = xS_1$ gives

$$\Pi_{\mathrm{A}} = \mathbb{E}_1\left[\mathbf{1}_{\{A_2=1\}}a + (1 - d\mathbf{1}_{\{A_2=1\}})bL_2^A\right] - xS_1. \tag{C.6}$$

Define the incremental gain from attack as

$$\Delta\Pi_{\text{liquid}} := \Pi_{\mathrm{A}}(A_2 = 1) - \Pi_{\mathrm{A}}(A_2 = 0) := a + (1 - d)\left[bL_2^A\right]_{A_2=1} - \left[bL_2^A\right]_{A_2=0}.$$

Attack is privately optimal whenever $\Delta\Pi_{\text{liquid}} > 0$. Define the deterrence cutoff $\bar{a}$ as the private benefit that makes the attacker indifferent:

$$\Delta\Pi_{\text{liquid}} = 0 \quad \Longleftrightarrow \quad \bar{a} = \left[bL_2^A\right]_{A_2=0} - (1 - d)\left[bL_2^A\right]_{A_2=1}. \tag{C.7}$$

C.4 Depositors' problem

Depositor participation keeps the same outside-option logic as equation (10), but expected backing is now pinned down by the compounding transition system. With par origination at date 1, depositors receive one LST per unit of ETH deposited, so $L_1^D = M_1^D$ and initial backing is one. The depositor's expected payoff is

$$\Pi_2^D = \mathbb{E}_1\left[(1 - d\mathbf{1}_{\{A_2=1\}})bL_2^D - M_1^D\right].$$

Since depositors do not receive direct rebase issuance, their claim quantity stays fixed: in equilibrium $L_2^D = L_1^D = M_1^D$. The marginal depositor is indifferent in equilibrium, so participation requires $\Pi_2^D \geq \mu\, M_1^D$, with equality giving

$$\mathbb{E}_1\left[(1 - d\mathbf{1}_{\{A_2=1\}})b\right] = 1 + \mu. \tag{C.8}$$

C.5 Participation condition of honest operators

Operator participation also keeps the same free-entry structure as in equations (12) and (14), subject to an additional fixed cost of entry $F > 0$. Another change is that, in the compounding model, the relevant payoff is summarized by the state-contingent claim value bL_2^O introduced in Section C.2, so backing dilution and token depreciation enter through a single continuation term.

Let L^O denote aggregate honest-operator LST holdings and let M^O denote aggregate ETH posted as bond capital. Let N denote the mass of identical honest operators, excluding the attacker. By symmetry, each operator holds

$$\bar{l}^O = \frac{L^O}{N} \quad \text{and} \quad \bar{m}^O = \frac{M^O}{N}.$$

Let $\bar{m}_1^O$ denote the bond posted by an individual operator at date 1 and let $\bar{l}_2^O$ denote its date-2 LST holdings after burn and rebase. The operator's expected payoff in ETH, net of the date-1 bond outlay, is

$$\Pi_2^O = \mathbb{E}_1\left[(1 - d\mathbf{1}_{\{A_2=1\}})b\bar{l}_2^O - \bar{m}_1^O\right]. \tag{C.9}$$

The operator therefore compares the expected value of the retained claim to the date-1 bond outlay, the fixed cost $F > 0$, and the outside option with net return μ per unit of posted bond capital. Participation requires

$$\Pi_2^O \geq F + \mu\, \bar{m}_1^O,$$

with equality in equilibrium.

Lemma 1 (Honest-operator free-entry condition)**.** *Honest-operator mass is given by*

$$N = \frac{1}{F}\left(\mathbb{E}_1\left[(1 - d\mathbf{1}_{\{A_2=1\}})bL_2^O\right] - (1+\mu)M_1^O\right). \tag{C.10}$$

Proof. Start from per-operator free entry, $\Pi_2^O = F + \mu \bar{m}_1^O$, and substitute equation (C.9). Multiplying both sides by N and using $M_1^O = N\bar{m}_1^O$ and $L_2^O = N\bar{l}_2^O$ gives

$$(1+\mu)M_1^O = \mathbb{E}_1\left[(1 - d\mathbf{1}_{\{A_2=1\}})bL_2^O\right] - NF.$$

Rearranging yields equation (C.10). ■

C.6 Equilibrium

With the scale accounting and transition map in place, equilibrium reduces to three familiar conditions: depositor participation (C.8), honest-operator free entry (C.10), and the attacker's best response. For any candidate bond ratio x, equation (C.3) maps depositor scale into (M_1^S, S_1, M_1^L), Section C.2 maps those date-1 quantities into state-contingent continuation values, and equilibrium requires those objects to be mutually consistent.

Lemma 2 (Equilibrium conditions, compounding model)**.** *Fix $x \in [0,1]$. An operating equilibrium is a triple (q, M_1^O, N) with $q \in [0,1]$, $M_1^O > 0$, and $N > 0$ such that:*

$$\mathbb{E}_1\left[(1 - d\mathbf{1}_{\{A_2=1\}})b\right] := (1-q)\left[b\right]_{A_2=0} + q(1-d)\left[b\right]_{A_2=1} := 1+\mu, \tag{C.11}$$

$$N = \frac{1}{F}\left(\mathbb{E}_1\left[(1 - d\mathbf{1}_{\{A_2=1\}})bL_2^O\right] - (1+\mu)M_1^O\right), \tag{C.12}$$

and one attacker condition:

(i) no-attack condition: $q = 0$ and $\Delta\Pi_{liquid} \le 0$;

(ii) attack condition: $q = 1$ and $\Delta\Pi_{liquid} \ge 0$;

(iii) mixed condition: $q \in (0,1)$ and

$$\Delta\Pi_{liquid} = 0 \quad \Longleftrightarrow \quad a = \bar{a}.$$

If no operating equilibrium exists, the equilibrium is shutdown with $M_1^O = 0$ and $N = 0$.

Proof. Equation (C.11) is depositor participation (C.8) written in state-contingent form. Equation (C.12) is exactly (C.10). The attacker conditions are the corresponding best-response conditions, with interior indifference given by (C.7). If no positive-scale triple satisfies these conditions, the residual outcome is shutdown. ■

Lemma 2 clarifies that depositor participation and operator entry both depend on state-contingent backing, while the attacker condition selects the relevant branch. Interior mixing is therefore possible only if those continuation objects move with q. If expected backing were fixed, pricing could support interior indifference only on a knife-edge set of parameters. In the compounding model, that dependence runs through the backing terms in equations (C.11) and (C.12), as shown below in Lemma 3.

Proposition C.2 (Equilibrium regimes by primitives)**.** *Fix $x \in (0,1)$. The following statements hold:*

(0) If $\left[b\right]_{A_2=0} = (1-d)\left[b\right]_{A_2=1}$, then no mixed equilibrium with $q \in (0,1)$ is supported except at the knife-edge equalities

$$\left[b\right]_{A_2=0} = 1+\mu, \qquad a = \left[bL_2^A\right]_{A_2=0} - (1-d)\left[bL_2^A\right]_{A_2=1}.$$

(i) If there exists $M_1^O > 0$ and $N > 0$ satisfying equation (C.12) at $q = 0$ and

$$\left[b\right]_{A_2=0} = 1+\mu, \qquad a \le \left[bL_2^A\right]_{A_2=0} - (1-d)\left[bL_2^A\right]_{A_2=1},$$

then there exists a no-attack operating equilibrium with $q = 0$.

(ii) If there exists $M_1^O > 0$ and $N > 0$ satisfying equation (C.12) at $q = 1$ and

$$(1-d)\left[b\right]_{A_2=1} = 1+\mu, \qquad a \ge \left[bL_2^A\right]_{A_2=0} - (1-d)\left[bL_2^A\right]_{A_2=1},$$

then there exists an attack operating equilibrium with $q = 1$.

(iii) If there exists (q, M_1^O, N) with $q \in (0,1)$, $M_1^O > 0$, and $N > 0$ satisfying equations (C.11), (C.12), and

$$a = \left[bL_2^A\right]_{A_2=0} - (1-d)\left[bL_2^A\right]_{A_2=1},$$

then

$$q = \frac{\left[b\right]_{A_2=0} - (1+\mu)}{\left[b\right]_{A_2=0} - (1-d)\left[b\right]_{A_2=1}} \in (0,1).$$

Thus interior mixing requires the denominator to be nonzero and the implied q to lie strictly between zero and one.

(iv) If no triple (q, M_1^O, N) satisfies (i), (ii), or (iii), then the equilibrium is shutdown.

Proof. For each candidate $M_1^O > 0$, equations (C.3) and (C.4) determine the date-1 scale and return objects, and Section C.2 then converts those objects into the state-contingent terms entering equations (C.7) and (C.11). Statements (i) and (ii) are the boundary regimes $q = 0$ and $q = 1$ from Lemma 2, together with equation (C.12).

For (iii), with $q \in (0, 1)$, equation (C.11) gives

$$(1 - q)\big[b\big]_{A_2=0} + q(1 - d)\big[b\big]_{A_2=1} = 1 + \mu,$$

which rearranges to the displayed formula for q when $\big[b\big]_{A_2=0} \neq (1 - d)\big[b\big]_{A_2=1}$. The equality in a is exactly the interior attacker indifference condition from equation (C.7). Statement (0) follows immediately: if $\big[b\big]_{A_2=0} = (1-d)\big[b\big]_{A_2=1}$, the pricing equation cannot pin down an interior q unless $\big[b\big]_{A_2=0} = 1+\mu$, and interior attacker indifference also requires the displayed equality in a. Statement (iv) is the residual case. ■

C.7 Security implications of liquid staking

This subsection compares deterrence cutoffs in the attacker's private benefit a, evaluated at the no-attack state. By Lemma 4, $M^S_{1,q=0}$, $S_{1,q=0}$, and $\lambda_{1,q=0}$ are pinned down independently of x. Let $a^*_{\text{traditional}}$ and $a^*_{\text{liquid}}(x)$ denote the traditional and liquid-staking cutoffs.

Corollary C.1 (Liquid staking weakens deterrence under limited bonding)**.** *If $x \leq \gamma$, then the endogenous deterrence cutoff for liquid staking is weakly smaller than the corresponding cutoff for traditional staking:*

$$a^*_{liquid}(x) \leq a^*_{traditional}.$$

Proof. Under traditional staking, the attacker stakes $S_{1,q=0}$ ETH directly without a pool or protocol fee. The deterrence cutoff is

$$a^*_{\text{traditional}} := S_{1,q=0}(\lambda_{1,q=0} + d + (1 - d)\gamma).$$

Write the gap as $a^*_{\text{traditional}} - a^*_{\text{liquid}}(x) = G_0 - (1 - d)G_1$, where

$$G_0 = (1 + \lambda_{1,q=0})S_{1,q=0} - \big[bL_2^A\big]_{A_2=0}, \qquad G_1 = (1 - \gamma)S_{1,q=0} - \big[bL_2^A\big]_{A_2=1},$$

with all date-1 objects evaluated at the no-attack equilibrium of Lemma 4. Here G_0 is the no-attack continuation gap between traditional and liquid staking, while G_1 is the

analogous gap in the attack state. Under no attack, the compounding mechanism gives $\left[b\right]_{A_2=0} = 1 + (1-\tau)\lambda_{1,q=0}$, and the attacker receives operator-fee rebase in proportion to its share of operator LSTs. This yields $\left[bL_2^A\right]_{A_2=0} := (1+(1-\tau)\lambda_{1,q=0})\, xS_{1,q=0} + \tau_N \lambda_{1,q=0}\, S_{1,q=0}$, so

$$G_0 = S_{1,q=0}\left[(1-x) + \lambda_{1,q=0}\big((1-x)(1-\tau_N) + \tau_P x\big)\right] \geq 0.$$

Under attack, the attacker's date-2 claim satisfies $\left[bL_2^A\right]_{A_2=1} \geq 0$ because both backing and holdings are nonnegative. Therefore $G_1 \leq (1-\gamma)S_{1,q=0}$ and $(1-d)G_1 \leq (1-d)(1-\gamma)S_{1,q=0}$. Combining them, we get

$$a^*_{\text{traditional}} - a^*_{\text{liquid}}(x) \geq S_{1,q=0}\left[(\gamma - x) + d(1-\gamma) + \lambda_{1,q=0}\big((1-x)(1-\tau_N) + \tau_P x\big)\right].$$

When $x \leq \gamma$, every term on the right is nonnegative. ■

For $x \leq \gamma$, liquid staking still weakens deterrence because the attacker internalizes less of the attack cost than under traditional staking. In the compounding model, that wedge is carried by the attacker's state-contingent retained claim rather than by a date-1 discounted sale price.

For $x > \gamma$, the ranking can reverse if the conservative burn rule over-penalizes the attacker relative to the neutral conversion benchmark, thereby creating an additional deterrence channel. In that region, there is no single closed-form cutoff in x alone: whether liquid staking is more or less deterrent than traditional staking depends on the full parameter configuration through the burn-rebase mapping.

Lemma 3 (Benchmark with exogenous expected backing)**.** *Consider a benchmark in which expected date-2 backing value is exogenously fixed at* $\mathbb{E}_1\left[(1 - d\mathbf{1}_{\{A_2=1\}})b\right] = 1 + \mu$ *and does not respond to information about attack intent. Then no mixed equilibrium with* $q \in (0,1)$ *exists, except possibly on a knife-edge set of parameters.*

If there exists (M_1^O, N) *with* $M_1^O > 0$ *and* $N > 0$ *satisfying operator participation and* $\Delta\Pi_{liquid} \leq 0$*, then there exists an operating equilibrium with* $q = 0$*. Otherwise, if there exists* (M_1^O, N) *with* $M_1^O > 0$ *and* $N > 0$ *satisfying operator participation and* $\Delta\Pi_{liquid} \geq 0$*, then there exists an operating equilibrium with* $q = 1$*.*

Proof. With expected backing value fixed at $1 + \mu$, the attacker's incremental gain $\Delta\Pi_{\text{liquid}}$ does not vary with q because the backing outcomes $\left[b\right]_{A_2=0}$ and $\left[b\right]_{A_2=1}$ are fixed. Therefore the indifference condition $\Delta\Pi_{\text{liquid}} = 0$ can hold only on a knife-edge set of parameters, so a mixed equilibrium with $q \in (0,1)$ cannot be supported in general.

Under fixed expected backing: if $\Delta\Pi_{\text{liquid}} \leq 0$ the attacker weakly prefers not attacking, giving $q = 0$; if $\Delta\Pi_{\text{liquid}} \geq 0$ the attacker weakly prefers attacking, giving $q = 1$. ■

This result carries over from the discount model: fixed expected backing removes the feedback needed to support interior mixing, so equilibria collapse to boundary regimes except at knife-edge parameter values.

C.8 Equilibrium properties

This subsection collects the comparative-static results used later in the protocol-design problem. The two questions are whether no-attack scale depends on x, and how positive attack risk feeds back into scale once operator participation is endogenous.

Lemma 4 (No-attack scale and bond-ratio independence)**.** *Suppose there exists an operating equilibrium with $q = 0$ and $M_1^S > 0$. Then*

$$(1-\tau)\lambda_1 = \mu,$$

and total staked ETH is uniquely pinned down by

$$M_{1,q=0}^S = \left(\frac{(1-\tau)k}{\mu}\right)^2,$$

with $\lambda_{1,q=0} = \mu/(1-\tau)$, $S_{1,q=0} = fM_{1,q=0}^S$, and $M_{1,q=0}^L = M_{1,q=0}^S - \alpha$, all independent of x. The component quantities $M_{1,q=0}^O = x\left[(1-f)M_{1,q=0}^S - \alpha\right]$ and $M_{1,q=0}^D = (1-x)M_{1,q=0}^L$ do depend on x.

Proof. Under the compounding mechanism in Section C.2, consider the no-attack state $A_2 = 0$. In step (i), all operators generate rewards and there is no slashing, so $\Delta M_2^L = \lambda_1 M_1^L$ and $M_2^L = (1+\lambda_1)M_1^L$. In step (ii), no penalty is applied and no LSTs are burned, so $\tilde{L}_2 = L_1$. In step (iii), since $\Delta M_2^L > 0$, the total issuance is $I_2 = \frac{\tau\lambda_1}{1+(1-\tau)\lambda_1}L_1$, so $L_2 = \frac{1+\lambda_1}{1+(1-\tau)\lambda_1}L_1$. Using $M_1^L/L_1 = 1$ at origination, date-2 backing becomes $\left[b\right]_{A_2=0} = M_2^L/L_2 = 1+(1-\tau)\lambda_1$. When $q = 0$, this is the only state, so depositor participation equation (C.8) gives

$$1 + (1-\tau)\lambda_1 = 1 + \mu,$$

hence $(1-\tau)\lambda_1 = \mu$. Substituting $\lambda_1 = k/\sqrt{M_1^S}$ from equation (C.4) yields $M_{1,q=0}^S = \left((1-\tau)k/\mu\right)^2$. The aggregate accounting in equation (C.3) then gives $S_{1,q=0} = fM_{1,q=0}^S$ and

$M^L_{1,q=0} = M^S_{1,q=0} - \alpha$, none of which involve x. The bond-ratio condition $(M^O_1 + M^A_1)/M^L_1 = x$ and $M^A_1 = xS_1$ yield $M^O_{1,q=0} = xM^L_{1,q=0} - xS_{1,q=0} = x\big[(1-f)M^S_{1,q=0} - \alpha\big]$, and $M^D_{1,q=0} = (1-x)M^L_{1,q=0}$. ■

The key implication is that the no-attack branch pins down aggregate scale through depositor participation alone. Changing x redistributes that scale across depositor and operator balance sheets, but it does not change total scale.

Lemma 5 (Positive attack risk lowers scale)**.** *Suppose there exists an operating equilibrium with $q > 0$ and $M^S_1 > 0$. Then $M^S_1 \le M^S_{1,q=0}$, with strict inequality if $d > 0$ or $x < \gamma$.*

Proof. Depositor participation equation (C.11) requires

$$(1-q)\big[b\big]_{A_2=0} + q(1-d)\big[b\big]_{A_2=1} = 1 + \mu.$$

From the compounding mechanism, $\big[b\big]_{A_2=0} = 1 + (1-\tau)\lambda_1$. Substituting,

$$1 + (1-\tau)\lambda_1 - q\Big(\big[b\big]_{A_2=0} - (1-d)\big[b\big]_{A_2=1}\Big) = 1 + \mu,$$

so $(1-\tau)\lambda_1 = \mu + q\big(\big[b\big]_{A_2=0} - (1-d)\big[b\big]_{A_2=1}\big)$. When $d > 0$ or $x < \gamma$, $\big[b\big]_{A_2=0} > (1-d)\big[b\big]_{A_2=1}$, so $(1-\tau)\lambda_1 > \mu$. Since $\lambda_1 = k/\sqrt{M^S_1}$ is strictly decreasing in M^S_1, a higher λ_1 requires $M^S_1 < M^S_{1,q=0}$. Equality $M^S_1 = M^S_{1,q=0}$ holds only when both $d = 0$ and $x \ge \gamma$, so that $(1-d)\big[b\big]_{A_2=1} = \big[b\big]_{A_2=0}$. ■

Positive attack risk therefore acts like an endogenous wedge in depositor participation. To keep depositors indifferent, the system must offer a higher staking return, which requires lower total scale unless both the dilution and depreciation channels are shut down ($d = 0$ and $x \ge \gamma$).

Proposition C.3 (Bonding and honest scale under positive attack risk)**.** *Suppose there exists an operating equilibrium with $q > 0$ and $M^S_1 > 0$. The depositor participation surplus is*

$$(1-\tau)\lambda_1 - \mu - q\left(\big[b\big]_{A_2=0} - (1-d)\big[b\big]_{A_2=1}\right).$$

Equilibrium participation requires this surplus to equal zero. Changes in x therefore operate through three conceptually distinct channels:

1. *The socialized-slashing channel operates through attack-state backing $\big[b\big]_{A_2=1}$: more first-loss bonding coverage x reduces the losses socialized to remaining claims and therefore tends to increase equilibrium honest scale M^S_1.*

2. *The depreciation-exposure channel operates through the attack-state term* $(1-d)\left[b\right]_{A_2=1}$*: more continuation value remains exposed inside the attacked system and therefore tends to decrease equilibrium honest scale* M_1^S.

3. *The deterrence-feedback channel operates through equilibrium attack risk* q*: changes in* x *shift attacker incentives and can therefore raise or lower equilibrium honest scale* M_1^S *indirectly.*

Accordingly, along positive-attack equilibria the effect of x *on equilibrium honest scale* M_1^S *is not signed in general.*

Proof. Start from depositor participation:

$$(1-q)\left[b\right]_{A_2=0} + q(1-d)\left[b\right]_{A_2=1} = 1+\mu.$$

Under the compounding mechanism, the no-attack backing state is

$$\left[b\right]_{A_2=0} = 1 + (1-\tau)\lambda_1.$$

Substituting and rearranging gives

$$(1-\tau)\lambda_1 - \mu - q\left(\left[b\right]_{A_2=0} - (1-d)\left[b\right]_{A_2=1}\right) := 0.$$

This is the participation-surplus expression in the proposition. Since $\lambda_1 = k/\sqrt{M_1^S}$ is strictly decreasing in M_1^S, any change in x that lowers the wedge $\left[b\right]_{A_2=0} - (1-d)\left[b\right]_{A_2=1}$ raises surplus at a given M_1^S and therefore supports a higher equilibrium scale.

The socialized-slashing channel is the compounding-model analogue of the discount-model uncovered-loss term. Through step (ii) of Section C.2, higher bonding increases first-loss capacity in the attacker's posted bond and reduces the loss that must be absorbed by the remaining claims, which tends to raise $\left[b\right]_{A_2=1}$ and lower the wedge. The depreciation-exposure channel is embedded in the same state-contingent object: because claims after an attack are multiplied by $(1-d)$, any change in the burn-rebase transition that leaves more continuation value exposed inside the attacked system lowers $(1-d)\left[b\right]_{A_2=1}$ relative to the no-attack benchmark and therefore raises the wedge. Finally, the deterrence-feedback channel does not appear as a separate additive term because it operates through equilibrium

probability of attack q. A change in x alters the attacker cutoff

$$\left[bL_2^A\right]_{A_2=0} - (1-d)\left[bL_2^A\right]_{A_2=1},$$

which can change the equilibrium attack probability and thereby magnify or offset the direct effects already described. If the burn-rebase mapping makes higher bonding weaken effective punishment in some region, equilibrium attack risk can rise and lower participation surplus through the wedge term above. When $x \geq \gamma$, there is zero uncovered slashing, so the pure backing-dilution part of the socialized-slashing channel disappears; however, burn-rebase transitions and induced attack incentives can still make the effect of x on protocol scale non-monotone. The overall effect of x on equilibrium honest scale is therefore not signed in general. ■

Proposition C.3 is the compounding-model counterpart of Proposition 4. The message is the same, but the mechanism is different. Along a positive-attack branch, higher bonding can either support or depress honest participation, so the effect of x is not signed in general. In the discount model, bonding enters participation partly through the date-1 discount term and partly through equilibrium attack risk. Here those forces are bundled into the attack-state backing wedge $\left[b\right]_{A_2=0} - (1-d)\left[b\right]_{A_2=1}$ and into the burn-rebase effect on the attacker's retained continuation payoff. Those same feedbacks also leave room for potentially multiple equilibrium branches at any given bond ratio.

C.9 The protocol's problem

The protocol now chooses (x, τ_N, τ_P) taking the equilibrium system as given. The question is whether a fee-maximizing protocol wants to tolerate attack risk in order to preserve activity, or instead prefers policies that implement the no-attack branch. In the compounding model, that comparison works through state-contingent backing and rebase revenue.

The protocol treasury receives the share τ_P of positive net pool earnings through the rebase mechanism. Using step (iii) of Section C.2, date-2 treasury fee revenue is

$$\Pi_2^P := \mathbf{1}_{\{\Delta M_2^L > 0\}} \tau_P \Delta M_2^L,$$

so the protocol objective is expected fee revenue

$$\Pi^P := \mathbb{E}_1\left[\Pi_2^P\right] := \mathbb{E}_1\left[\mathbf{1}_{\{\Delta M_2^L > 0\}} \tau_P \Delta M_2^L\right].$$

Lemma 6 (No-attack protocol payoff and scale)**.** *In an operating no-attack equilibrium ($q = 0$), for $\tau_P \in (0, \tau]$, Π^P is strictly increasing in $M_1^S > 0$.*

Proof. If $q = 0$, then step (i) of Section C.2 gives $\Delta M_2^L = \lambda_1 M_1^L > 0$, so $\Pi_2^P = \tau_P \lambda_1 M_1^L$ and therefore $\Pi^P = \tau_P \lambda_1 M_1^L$. Using $\lambda_1 = k/\sqrt{M_1^S}$ from equation (C.4) and $M_1^L = M_1^S - \alpha$, the protocol payoff becomes

$$\Pi^P := \frac{\tau_P k (M_1^S - \alpha)}{\sqrt{M_1^S}}.$$

Differentiating with respect to M_1^S:

$$\frac{d\Pi^P}{dM_1^S} := \frac{\tau_P k (M_1^S + \alpha)}{2(M_1^S)^{3/2}} > 0.$$

■

Proposition C.4 (Profit-maximizing protocol implements no-attack regime)**.** *For given $\tau \in (0, 1]$, if some x supports an equilibrium with $q = 0$, then a profit-maximizing protocol chooses such an x.*

Proof. By Lemma 4, the no-attack equilibrium scale $M_{1,q=0}^S$ is independent of x, so the protocol payoff under $q = 0$ is the same for every x that supports it.

Fix any operating equilibrium with $q > 0$ at some x'. State by state, step (i) of Section C.2 implies $\Delta M_2^L \le \lambda_1 M_1^L$, hence

$$\mathbf{1}_{\{\Delta M_2^L > 0\}} \tau_P \Delta M_2^L \le \tau_P \lambda_1 M_1^L.$$

Taking expectations gives

$$\Pi^P \le \tau_P \lambda_1 M_1^L.$$

By Lemma 5, any such $q > 0$ equilibrium satisfies $M_1^S \le M_{1,q=0}^S$, with strict inequality when $d > 0$ or $x' < \gamma$. By Lemma 6, the no-attack payoff formula $\tau_P \lambda_1 M_1^L$ is strictly increasing in M_1^S. Therefore

$$\Pi^P(q > 0) \le \tau_P \lambda_1 (q > 0) M_1^L (q > 0) \le \tau_P \lambda_1 (q = 0) M_1^L (q = 0) = \Pi^P (q = 0).$$

Thus, the $q = 0$ outcome weakly dominates any $q > 0$ outcome under this objective. A profit-maximizing protocol therefore selects x to implement the no-attack regime whenever such x exists. ■

The revenue ranking is therefore clear: whenever the no-attack branch can be implemented, a fee-maximizing protocol prefers it. The remaining issue is implementability. Even if the protocol prefers the no-attack branch, can bonding alone select it uniquely? Let $M^S_{1,q=0} > 0$ denote date-1 equilibrium total staked ETH in an operating no-attack equilibrium, as pinned down in Lemma 4.

Proposition C.5 (Attack feasibility and the bond cutoff at $q = 0$)**.** *At the $q = 0$ operating candidate $M^S_1 = M^S_{1,q=0}$ satisfying the feasibility condition (C.5), the no-attack deterrence condition is*

$$a \le a^*_{liquid}(x) := \left[bL^A_2\right]_{A_2=0} - (1-d)\left[bL^A_2\right]_{A_2=1},$$

where both terms are evaluated at the no-attack equilibrium via the compounding mechanism in Section C.2. The quantities $M^S_{1,q=0}$, $S_{1,q=0} = fM^S_{1,q=0}$, and $\lambda_{1,q=0}$ are independent of x (Lemma 4). Therefore, at the $q = 0$ operating candidate, the dependence of $a^*_{liquid}(x)$ on x is entirely the dependence of $\left[bL^A_2\right]_{A_2=0} - (1-d)\left[bL^A_2\right]_{A_2=1}$ on the bond-to-pool ratio through the burn mechanism.

Proof. Using equation (C.3), $M^L_1 = M^S_1 - \alpha$ and $S_1 = fM^S_1$. Therefore $M^L_1 - S_1 = (1-f)M^S_1 - \alpha$. The feasibility condition (C.5) is equivalent to $M^S_1 \ge \alpha/(1-f)$, so $M^L_1 \ge S_1$.

For the deterrence claim, fix $q = 0$ and substitute the no-attack equilibrium from Lemma 4. The incremental attack gain is $\Delta\Pi_{\text{liquid}} = a + (1-d)\left[bL^A_2\right]_{A_2=1} - \left[bL^A_2\right]_{A_2=0}$, so $\Delta\Pi_{\text{liquid}} \le 0$ if and only if $a \le a^*_{\text{liquid}}(x)$. Since $M^S_{1,q=0}$, $S_{1,q=0}$, and $\lambda_{1,q=0}$ do not depend on x, the only channel through which x enters a^*_{liquid} is the bond-to-pool ratio x governing the burn mechanism in Section C.2. ■

This proposition separates scale from deterrence on the no-attack branch. At the candidate with $q = 0$, aggregate scale is already pinned down, so bonding affects deterrence only through loss allocation. In the compounding model, that channel is summarized by $\left[bL^A_2\right]_{A_2=0} - (1-d)\left[bL^A_2\right]_{A_2=1}$, namely the burn-rebase mapping from the bond ratio to the attacker's retained collateral claim.

The protocol problem therefore separates preference from implementability. Proposition C.4 shows that the protocol prefers an operating equilibrium with $q = 0$ whenever such an equilibrium is available, because fee revenue is highest on the no-attack branch. But that preference does not imply that bonding alone can implement a unique no-attack equilibrium.

Proposition C.3 shows why. Increasing x changes participation through both the backing wedge and induced attack incentives, so the compounding model can admit multiple self-

consistent branch values of (q, M_1^S). Even if the protocol sets $x \geq \gamma$ and thereby removes uncovered slashing from the direct backing-dilution channel, a low-scale branch with lower M_1^S and a lower required attack stake S_1 can still, in principle, sustain a boundary attack equilibrium through the burn-rebase mapping. Bonding policy is therefore not sufficient in general to uniformly implement a unique equilibrium with $q = 0$.

The comparison with the discount model is close in conclusion but different in mechanism. In the discount model, the main implementation difficulty comes from the interaction among operator participation, the date-1 discount, and attack incentives. In the compounding model, the same fixed-point problem is carried by contingent-state backing and the attacker's retained continuation claim after burn and rebase. The policy implication is unchanged: robust implementation of the no-attack regime generally requires complementary tools in addition to bonding, or sufficiently precise knowledge of the underlying parameters to identify a region in which no attack is uniquely sustained.

C.10 Attacker participation and the outside option of traditional staking

This final subsection adds a robustness margin in the spirit of Section B for the discount model. Before choosing attack or no attack, the attacker can decide whether to use the liquid-staking strategy at all. The question is whether allowing that opt-out changes the protocol ranking derived above.

Suppose the attacker can deploy the same capital in traditional staking at net return λ_1 per unit of ETH, and that this venue choice does not change aggregate date-1 stake M_1^S. In the compounding model, the relevant participating capital is the bond outlay xS_1. If the attacker uses the liquid-staking strategy, its continuation value is determined by the burn-rebase dynamics in Section C.2; if it opts out, it receives $\lambda_1 \cdot xS_1$.

When the attacker participates, its expected payoff net of the initial bond outlay is given by equation (C.6). Conditional on attack probability q, this becomes

$$\mathbb{E}_1[\Pi_A \mid q] = (1-q)\left[bL_2^A\right]_{A_2=0} + q\left(a + (1-d)\left[bL_2^A\right]_{A_2=1}\right) - xS_1.$$

The participation constraint requires this to weakly exceed the outside option:

$$\mathbb{E}_1[\Pi_A \mid q] \;\geq\; \lambda_1 \cdot xS_1. \tag{C.13}$$

Lemma 7 (Attacker participation condition)**.** *The attacker's expected payoff conditional on attack probability q is*

$$\mathbb{E}_1[\Pi_A \mid q] = \left[bL_2^A\right]_{A_2=0} + q\,\Delta\Pi_{liquid} - xS_1,$$

where $\Delta\Pi_{liquid} = a + (1-d)\left[bL_2^A\right]_{A_2=1} - \left[bL_2^A\right]_{A_2=0}$ is the incremental attack gain. The participation constraint equation (C.13) *holds if and only if*

$$\left[bL_2^A\right]_{A_2=0} + q\,\Delta\Pi_{liquid} \;\geq\; (1+\lambda_1)\,xS_1. \tag{C.14}$$

Proof. Expanding equation (C.6) over the two branches gives

$$\begin{aligned}\mathbb{E}_1[\Pi_A \mid q] &= (1-q)\left[bL_2^A\right]_{A_2=0} + q\left(a + (1-d)\left[bL_2^A\right]_{A_2=1}\right) - xS_1 \\ &= \left[bL_2^A\right]_{A_2=0} + q\,\Delta\Pi_{\text{liquid}} - xS_1.\end{aligned}$$

Substituting into equation (C.13) gives the stated condition. ■

The participation condition (C.14) is imposed jointly with the equilibrium system. It is an additional consistency condition rather than a replacement for the attacker's best response.

Proposition C.6 (No-attack optimality under attacker non-participation)**.** *Suppose the optimal protocol policy implements an operating equilibrium with $q = 0$. Then that same policy remains optimal in the extended problem in which the attacker can opt out and stake honestly in traditional staking.*

Proof. Fix the baseline-optimal policy and its operating equilibrium with $q = 0$, with date-1 equilibrium total staked ETH $M^S_{1,q=0}$ and required attacked stake $S_{1,q=0} = fM^S_{1,q=0}$. The attacker's payoff from participating and not attacking is $\left[bL_2^A\right]_{A_2=0} - xS_{1,q=0}$. The outside option yields $\lambda_{1,q=0} \cdot xS_{1,q=0}$.

Under $q = 0$, no slashing penalty is realized. By the maintained assumption in this subsection, honest participation in the liquid-staking strategy and the outside option of traditional staking deliver the same per-unit continuation return $\lambda_{1,q=0}$ on the capital $xS_{1,q=0}$. Therefore

$$\left[bL_2^A\right]_{A_2=0} - xS_{1,q=0} = \lambda_{1,q=0} \cdot xS_{1,q=0}.$$

So the participation constraint (C.13) is satisfied with equality. There exists an equilibrium in which the attacker participates but does not attack, and the protocol payoff under the $q = 0$ policy is unchanged. Since the extended problem only adds the attacker's non-participation

option, a policy that is optimal in the baseline and implements $q = 0$ remains feasible and optimal. ■

Allowing the attacker to opt out therefore adds a participation constraint, but it does not overturn the no-attack policy ranking. Under $q = 0$, participation and opt-out give the same return on $xS_{1,q=0}$, so a policy that implements $q = 0$ remains optimal.

References

Jermann, Urban J., "A Macro Finance Model for Proof-of-Stake Ethereum," January 2023.

Jermann, Urban J., "Optimal Issuance for Proof-of-Stake Blockchains," September 2024.