

An Impossibility Theorem for Price-Based Risk Constraints

Francesco Nicolai*

BI Norwegian Business School

February 2026 – Latest Version

July 2024 – First Version

Abstract

Price-based constraints map sampled transaction prices into a risk statistic and then into a binding requirement, such as posted margin, a haircut, a leverage limit, or a mandated portfolio reallocation. We prove a constructive local impossibility theorem: no price-based, risk-sensitive rule can jointly achieve three desirable properties, namely (i) risk sensitivity, (ii) liquidity continuity, (iii) and round-trip manipulation-proofness, in any reachable binding state in which amplification is sufficiently strong. The proof constructs a rule-induced trigger-and-reverse deviation: an arbitrarily small trade tilts the sampled statistic, tightens a binding requirement at first order, induces predictable forced liquidation, and is profitably reversed. Unlike classic manipulation that relies on large trades to move the price level, the incentive here runs through first-order effects of sampled prints on the rulebook and the resulting forced flow; it does not require large trades or large price moves, and it does not rely on making a slack constraint bind. The result holds with continuous trading between discrete measurement and reset times, and it survives strategic responses by constrained traders.

Vulnerability is state dependent: marginal sensitivity is highest when realized risk is low, so tranquil states can be most exposed to manipulation. Quantitative relevance is summarized by a local amplification factor, the product of pass-through elasticities along the loop from prices to the statistic, to margin requirements, to forced sales, and back to prices via impact. With portfolio margining and cross-impact, trigger and liquidation directions can differ, generating cross-contract manipulation and spillovers. The design implication is immediate: implementability requires limiting short-horizon pass-through, rationalizing anti-procyclicality tools such as smoothing and state-contingent buffers.

Keywords: collateral design; margins and haircuts; secured funding constraints; endogenous risk measurement; central clearing.

JEL Classification: G21, G28, G12, D82, D62.

*Department of Finance, BI Norwegian Business School, Nydalsveien 37, 0484, Oslo, Norway, email: francesco.nicolai@bi.no.

1 Introduction

Many contracts and market designs update binding risk constraints from recent transaction prices, including margins and haircuts, internal VaR or expected shortfall (ES) limits, rule-based leverage targets, and portfolio rebalancing rules. When these constraints bind, the measured input is not passive: sampled transaction prices determine next-period balance-sheet capacity, so trading that affects the sampled marks can mechanically affect future requirements and the resulting balance-sheet pressure. In such settings, the update rule itself becomes part of the trading environment.

In this paper, we prove a constructive local impossibility theorem for price-based, risk-sensitive constraint updates. To fix ideas we write $M_{t+1} = g(\Gamma_t)$ and refer to M_{t+1} as margin, but the theorem applies to any binding requirement computed mechanically from sampled transaction prices. Consider any mechanism that sets next-period requirements from a risk statistic (Γ_t) computed from sampled transaction prices, $M_{t+1} = g(\Gamma_t)$. When M_{t+1} binds, it induces forced liquidation with price impact. In such states, no update rule can simultaneously satisfy: (i) *local risk sensitivity*, meaning that posted requirements increase at first order when the measured risk statistic rises; (ii) *liquidity continuity*, meaning that prices and aggregate positions vary continuously under small shocks; and (iii) *round-trip manipulation-proofness*, meaning that no admissible finite-horizon round trip earns strictly positive expected profit net of impact and any modeled quadratic inventory cost through its effect on the next update. If (i) and (ii) hold, then whenever amplification is sufficiently strong, round-trip manipulation-proofness fails in reachable binding states (Theorem 1).¹ Thus the impossibility is general ex-ante: if a proposed rule admits reachability of such a binding high-amplification state, no designer can make the rule satisfy all three properties there.

The result is important because each property, taken in isolation, is a standard objective in collateral design and risk management, and much of the literature effectively imposes one or two of them as maintained assumptions. Risk sensitivity aligns requirements with measured exposure, supporting coverage and incentives (Biais et al., 2016; Wang et al., 2022). Liquidity continuity is a market-functioning objective in environments where binding constraints interact with price impact; without continuity, small shocks can trigger discontinuous liquidity dry-ups and fire-sale dynamics that are widely viewed as destabilizing and welfare-reducing (Brunnermeier and Pedersen, 2009; Shleifer and Vishny, 2011). Manipulation-proofness is an incentive-compatibility requirement: in price-impact settings, profitable round trips correspond to canonical notions of manipulation or dynamic arbitrage and are excluded by standard no-manipulation conditions (Allen and Gale, 1992; Huberman and Stanzl, 2004). Our theorem shows that these three desiderata cannot generally and universally be satisfied simultaneously.

With exogenous constraints, a sufficiently small round trip should lose to price impact, and an infinitesimal price perturbation should not matter for tomorrow’s feasibility. With price-based constraints, by contrast, sampled transaction prices are inputs to next-period balance-sheet capacity through the update of a binding requirement. In binding states, a tiny, temporary distortion of a sampled mark moves the measured risk statistic and therefore moves the next posted requirement at

¹We avoid the label “margin trilemma” because in policy discussions the “margin trilemma” for CCP initial margin is an objective trade-off: improving statistical coverage typically requires either higher average margin levels or more reactive (more procyclical) margin updates. Our result is about a different set of properties.

first order. That tighter requirement forces constrained investors to cut positions; their forced sales (or purchases) are price-impactful and mechanically propagate the initial mark distortion into subsequent prices. When pass-through from marks into requirements is steep enough, the rule-induced flows can be large relative to the manipulator’s own impact costs, so a trigger-and-reverse strategy can be profitable even though fundamentals have no predictable drift. This restriction is extremely general. It applies whenever a binding feasibility constraint is updated mechanically from a statistic of past transaction prices in a market that is not infinitely liquid. In such environments, even arbitrarily small perturbations of sampled marks can shift tomorrow’s constraint at first order and generate predictable, rule-driven order flow.²

Several clarifications delimit the scope of the theorem and address common misreadings. First, the result is local and it is about binding regimes. It applies in neighborhoods of reachable states in which the posted requirement already binds and a marginal tightening generates marginal forced deleveraging with nonzero price pressure. In those states, the manipulator does not need a large trade or a large price move: any profitable trigger is itself small and remains within a local feasibility bound (Theorem 1); the mechanism operates through first-order pass-through from a sampled mark into tomorrow’s binding requirement.

Second, constrained investors are allowed to be forward-looking: if they anticipate manipulation, the resulting reduction in forced-flow intensity is an equilibrium response.³ Section 4.3 microfounds strategic constrained traders: anticipation can dampen forced flows, but it cannot restore joint feasibility in any state in which the requirement still binds. The tension disappears only in degenerate regimes: either constrained participation collapses exactly in stress so that forced flow vanishes, or the mechanism is avoided by keeping requirements slack precisely when risk-sensitive requirements are meant to tighten. In both cases, the feedback is eliminated only by shutting down risk-sensitivity where it matters most.

Third, allowing continuous trading between discrete update times does not remove the mechanism. The phenomenon is pinned to the reset: in binding states, a small perturbation of a sampled mark at the update moves the next posted requirement at first order, which creates predictable post-update balance-sheet pressure and order flow. Section 4.1 shows that the same two-step, trigger-and-reverse logic can be executed around update times even when trading is continuous between them. More generally, any implementation based on discrete marks, auctions, jumps, kinks, or other non-smooth overlays retains directional sensitivity in binding states, so the same concern survives outside the knife-edge case of a fully smooth, continuously updated constraint.

Fourth, the mechanism is not classic manipulation in the sense of persistent mispricing, spoofing, corners, or settlement-print distortions. The predictable component is rule-induced: sampled transaction prices mechanically map into next-period requirements, and when requirements bind that

²Nicolai (2026), for instance, treats the rebalancing rule as given and converts the amplification condition into a liquidity-based capacity bound. The paper computes the maximum aggregate scale of price-insensitive (inelastic) AUM, such as volatility-managed strategies that rebalance mechanically, that a market can absorb without admitting a profitable trigger-and-reverse manipulation. The bound is stated in terms of maximum AUM as a function of average daily volume and an impact calibration, delivering a concrete ceiling on the size of the inelastic sector in a given market.

³Consistent with this equilibrium-response view, transaction-level evidence from cleared repo shows that higher CCP haircuts relative to OTC induce venue substitution and adverse selection, with safer borrowers shifting away from the CCP and CCP borrower composition deteriorating (Chebotarev, 2025).

mapping generates predictable post-update forced flow and associated price pressure even when the unaffected price is a martingale and the underlying impact block is manipulation-free absent feedback (in the [Huberman and Stanzl \(2004\)](#) sense).⁴

A further implication, often obscured by standard risk diagnostics, is that local manipulability can be highest in tranquil states. Even when the posted requirement is expressed as VaR or ES, it is typically pinned down by an estimated volatility (possibly filtered or scaled) computed from a window of recent returns. For example, with $\hat{\sigma}_t = \sqrt{RV_t/(m-1)}$ and $RV_t = \sum_{j=0}^{m-1} R_{t-j}^2$, holding the most recent return R_t fixed, the local sensitivity to the last sampled mark scales as $R_t/\sqrt{(m-1)RV_t}$; hence the same small mark distortion moves the risk input more after a quiet spell (low RV_t). Tranquil states can therefore be most vulnerable to trigger-and-reverse manipulation despite appearing safest ex ante, a rule-based analogue of the “paradox of financial instability” ([Borio and Drehmann, 2009](#)). This differs from the leverage-cycle mechanism, where low measured risk relaxes margins or haircuts and endogenously expands equilibrium leverage and positions; here, even without balance-sheet expansion, low realized volatility makes the input most sensitive to a single sampled mark, so small price distortions shift the next requirement most in tranquil states ([Danielsson et al., 2004](#); [Brunnermeier and Pedersen, 2009](#); [Adrian and Shin, 2010, 2014](#); [Geanakoplos, 2010](#); [Brunnermeier and Sannikov, 2014](#)).

Theorem 1 is proved by an explicit two-step trigger-and-reverse strategy. The proof delivers a sharp diagnostic. Profitability is governed by a simple amplification chain: how strongly the rule-book maps marks into requirements, how strongly requirements map into forced selling, and how strongly forced selling moves prices. These components are, in principle, calibratable from data on margin methodology, liquidation behavior, and market impact. Read as a design restriction, the diagnostic implies an upper bound on short-horizon pass-through from price-based inputs into posted requirements if the mechanism is to be robust to trigger-and-reverse manipulation. We use this restriction to discipline a margin-design problem. The optimal schedules feature pooling and deliberately muted adjustment regions, providing a structural rationale for standard anti-procyclicality tools such as limited pass-through, smoothing, and state-contingent buffers.

Finally, we extend the logic to portfolio margining with cross-impact. Portfolio systems map a vector of marks into a single scalar charge, but a binding call is met through a vector liquidation response that can be concentrated in contracts different from those that most efficiently move the portfolio risk statistic. This decoupling creates a trigger-versus-harvest wedge: a trader can tilt the portfolio risk number using one set of instruments while positioning in the instruments that will be forced to unwind when the call tightens, so incentives propagate across contracts and generate spillovers. We show that the existence and strength of this cross-contract channel is governed by an interpretable alignment term between the local risk-increasing direction in mark space and the liquidation-driven price-pressure direction implied by cross-impact; when this alignment is strong enough, profitable cross-asset round trips exist for arbitrarily small admissible triggers. A sharp

⁴In these regards, [Nicolai \(2026\)](#) constructs an explicit example in which (i) the unaffected price is a martingale and (ii) the impact specification is dynamically manipulation-free for unconstrained trading in the [Huberman and Stanzl \(2004\)](#) sense, so profitable round-trip price manipulation is ruled out absent feedback. Nevertheless, once a disclosed price-based constraint maps sampled transaction prices into next-period requirements or mandated positions, a trader can profit via a trigger-and-reverse deviation that exploits the rule-induced, predictable post-update forced flow, rather than any persistent mispricing.

empirical implication follows: the contracts that exhibit trading pressure inside the sampling window (the trigger leg) need not be the contracts that experience post-update illiquidity and forced-flow pressure (the liquidation leg), so portfolio margining can relocate stress across legs of the cleared complex.

1.1 Contribution and related literature

The paper sits at the intersection of three literatures that are usually studied separately. First, work on amplification through binding constraints and price impact studies forced sales, liquidity spirals, and endogenous leverage, and sometimes endogenizes margins or haircuts as functions of volatility or liquidity; it typically does not model a rulebook in which a short sampling window of transaction prices mechanically determines the next posted requirement in a way that gives a single trader first-order control over the update (Kiyotaki and Moore, 1997; Geanakoplos, 2010; Brunnermeier and Pedersen, 2009; Brunnermeier and Sannikov, 2014; He and Krishnamurthy, 2013; Gromb and Vayanos, 2002; Gârleanu and Pedersen, 2011). Second, the procyclicality and risk-regulation literature often treats the risk input as exogenous or statistically specified (for example, volatility dynamics), so the incentive problem created by estimating the input from tradable marks is not isolated as a design constraint, even though related work emphasizes feedback from constraints into prices (Basak and Shapiro, 2001; Danielsson et al., 2004; Adrian and Shin, 2014; Glasserman and Wu, 2018). Third, in classic no-manipulation and no-dynamic-arbitrage models with market impact and exogenous constraints, sufficiently small round trips lose to impact and cannot generate strictly positive expected profits (Allen and Gale, 1992; Jarrow, 1992; Huberman and Stanzl, 2004; Gatheral, 2010). Price-based constraints add a missing ingredient: sampled transaction prices enter next-period binding feasibility through an endogenous update $M_{t+1} = g(\Gamma_t)$, so in binding states an arbitrarily small perturbation of a sampled mark shifts the next requirement at first order and creates rule-induced predictable post-update order flow.

Our contribution is to take the update rule itself as the object of analysis and to characterize its design frontier. First, we prove a constructive local impossibility: near reachable binding states, no price-based, risk-sensitive update rule can jointly satisfy risk sensitivity, liquidity continuity, and round-trip manipulation-proofness (Theorem 1). Second, the deviation gain collapses to a single local amplification chain, yielding a sharp vulnerability diagnostic and, read as a restriction, an implied cap on short-horizon pass-through from price-based inputs into posted requirements (Section 3). Third, under portfolio margining with cross-impact, portfolio risk aggregation and portfolio liquidation need not align; this wedge generates cross-contract manipulation and spillovers (Section 5). Fourth, imposing monotonicity together with the implied slope restriction delivers optimal pooling-plus-adjustment schedules with bounded slopes; discontinuities do not eliminate incentives but relocate them to threshold-crossing deviations (Section 6). The margin-design exercise uses a parsimonious reduced-form objective aligned with PFMI-style coverage and stability trade-offs to isolate how the implementability restriction shapes pooling (Committee on Payments and Market Infrastructures and International Organization of Securities Commissions, 2012).

This paper provides the foundational impossibility result for price-based constraints: it characterizes the universal incentive restriction created by using tradable marks as inputs to binding feasibility

in markets with price impact. Nicolai (2026) applies the same logic in a different environment. Taking a deterministic rulebook g as given, he converts the local amplification condition into a viability screen and a liquidity-scaled capacity bound for mechanically rebalanced, price-insensitive capital (for example volatility-managed portfolios). In this sense, our theorem isolates the implementability constraint, while Nicolai (2026) shows how to deploy the diagnostic to quantify admissible scale in other settings. Concrete non-CCP instances include internal bank VaR/ES and leverage-limit rulebooks pinned to recent marks, secured-funding haircuts in repo or prime brokerage tied to volatility and liquidity metrics, volatility-control indices that mandate mechanical reallocations, and risk-parity or volatility-targeting mandates that scale exposure mechanically with estimated risk; more generally, any mechanical investment mandate in which exposure is a deterministic function of recent price realizations, realized volatility, drawdowns, or other price-based statistics fits the same mapping.

The design implications speak directly to clearing and margin policy. In clearing economics, central clearing manages counterparty default risk through margining and default resources, and netting sets and clearing architecture shape exposures and collateral demand (Duffie and Zhu, 2011; Menkveld and Vuillemeij, 2021; Cont and Kokholm, 2014; Loon and Zhong, 2014; Duffie et al., 2015). Transaction-level evidence from repo markets shows that haircut setting affects participation and composition: bilateral haircuts vary with counterparty type and relationships, while uniform CCP haircuts can induce venue substitution and adverse selection when they rise relative to OTC (Julliard et al., 2024; Chebotarev, 2025). Policy frameworks emphasize both coverage and stability of initial margin and document anti-procyclicality tools⁵; empirical and policy work shows that margin calls generate large, state-dependent liquidity demands in stress (Murphy et al., 2014, 2016; Aldasoro et al., 2023; King et al., 2023). Our mechanism adds an incentive constraint to these discussions: once the risk input is computed from tradable marks and requirements bind, short-horizon pass-through becomes a market-design object with an incentive-driven upper bound, providing a structural rationale for slope limits, smoothing, and buffer-type overlays. This perspective also connects to work on robustness and incentives in risk measurement and margin setting, and to analyses of CCP governance and externalities in stress (Cont et al., 2010; Huang and Takáts, 2024; Cont and Ghamami, 2025; Pirrong, 2011, 2014). A particularly clean contracting benchmark is Kuong and Maurin (2024), who endogenize key layers of the default waterfall by modeling the CCP as an agent. Our contribution is complementary: taking the institutional role of collateral and default resources as given, we isolate an additional implementability constraint that arises whenever the posted requirement is computed from tradable marks and binds, so the mapping from marks to requirements becomes part of the trading environment.

Roadmap. Section 2 sets up a canonical cleared-market model with $M_{t+1} = g(\Gamma_t)$ to fix ideas; the analysis and results apply to any price-based constraint that maps sampled transaction prices into a binding requirement. The main text uses linear impact, realized variance, and an explicit margin-

⁵See Committee on Payments and Market Infrastructures and International Organization of Securities Commissions (2012); Basel Committee on Banking Supervision (2022); Basel Committee on Banking Supervision and Committee on Payments and Market Infrastructures and Board of the International Organization of Securities Commissions (2022); European Union (2013); European Securities and Markets Authority (2018); European Systemic Risk Board (2020); Bank of England (2021); European Central Bank (2023).

elastic demand rule for transparency; the general case is proved in Appendix A. Section 3 formalizes the three desiderata, derives the round-trip strategy, and proves the local non-joint-feasibility theorem; it then rewrites the profit lower bound in terms of the loop gain to obtain the slope cap and a magnitude screen. Section 4 studies robustness and implementation, including continuous trading between discrete update times, alternative risk inputs (parametric VaR/ES, scenario maxima and quantiles), and implementation overlays such as trimming, thresholds, rounding, floors, and step-wise add-ons. Section 5 extends to portfolio margining with cross-impact and forced liquidation across contracts and establishes a corresponding cross-contract impossibility result. Section 6 solves the constrained margin-design problem and characterizes pooling-plus-adjustment rules.

2 Model

We formalize the environment of the introduction. A margin designer (e.g., a CCP) computes a risk statistic from transaction prices and updates initial margin mechanically. Near reachable binding states, a marginal change in M_{t+1} induces forced liquidation with a nonzero local slope; the resulting rulebook mapping can therefore generate predictable, liquidation-driven price pressure that a finite-horizon trigger-and-reverse strategy can monetize. The analysis is local: it does not rely on large trades, large price moves, or on making a slack constraint become binding. For transparency, the main text adopts three special cases: linear price impact in net order flow, realized variance from a fixed sampling window as the risk input, and a target-with-cap liquidation policy that makes the position elasticity with respect to margin explicit. These are expositional. Appendix A replaces realized variance with a generic price-based risk functional, allows a general local impact map, and permits broader liquidation responses. Whenever the main proof uses special-case structure, we flag it.

2.1 Primitives and timing

Time is discrete, $t = 0, 1, 2, \dots$. A single cleared contract is traded and marked to market each period. Fundamentals $V_t \in \mathbb{R}$ are exogenous, adapted to the public filtration $\{\mathcal{F}_t\}$, and satisfy

$$\mathbb{E}[V_{t+1} \mid \mathcal{F}_t] = V_t. \quad (1)$$

Assumption (1) removes predictable drift from the strategic trader's expected profit. Fix $m \geq 2$. Let P_t be the period- t transaction price and M_t the per-unit initial margin applied in period t . The public state at the start of t is

$$S_t = (V_t, P_{t-m}, \dots, P_{t-1}, M_t).$$

A continuum of constrained traders $i \in [0, 1]$ choose end-of-period positions $x_t(i)$ subject to M_t . A strategic trader chooses $a_t \in [-\bar{a}, \bar{a}]$, $\bar{a} > 0$, and his inventory evolves as $y_t = y_{t-1} + a_t$ with $y_{-1} = 0$. Competitive liquidity suppliers clear net order flow and P_t is realized by the price formation rule specified below. The margin designer computes a price-based input from the realized window and

posts next-period margin,

$$\Gamma_t = \Gamma(P_{t-m+1}, \dots, P_t), \quad M_{t+1} = g(\Gamma_t).$$

At $t + 1$, M_{t+1} applies. Breaching accounts adjust immediately through forced liquidation, and this order flow enters the same price formation rule in period $t + 1$. Constrained traders choose $x_t(i)$ as a function of S_t and do not condition on P_t , which is realized only after orders clear. They may condition on P_t from $t + 1$ onward since $P_t \in S_{t+1}$. Allowing within-window conditioning or re-optimization only changes the effective local elasticity of forced flow with respect to M_{t+1} ; the mechanism applies whenever that effective slope is strictly positive.

2.2 Price formation

Let $X_t = \int_0^1 x_t(i) di$ be the aggregate constrained position at the end of period t , and let $\Delta X_t = X_t - X_{t-1}$ be constrained traders' net order flow in period t . Total net order flow is

$$Q_t = \Delta X_t + a_t.$$

Transaction prices follow linear price impact with slope $\alpha > 0$:

$$P_t = V_t + \alpha Q_t = V_t + \alpha(\Delta X_t + a_t). \quad (2)$$

Thus α is the local price response to order flow, including forced-liquidation flow. Appendix A.4 allows a general local impact map; for the main results only the local sensitivity of P_t to Q_t around the benchmark path matters.

2.3 Margin update rule and price-based risk input

The margin designer (e.g., a CCP) sets next-period per-unit initial margin by

$$M_{t+1} = g(\Gamma_t), \quad (3)$$

where $g : \mathbb{R}_+ \rightarrow (0, \infty)$ is public and increasing. In the baseline model the risk input is realized variance computed from transaction prices over the m -period window:

$$\Gamma_t = RV_t = \sum_{j=0}^{m-1} (P_{t-j} - P_{t-j-1})^2. \quad (4)$$

Since transaction prices satisfy (2), trading within the window moves the sampled marks and therefore shifts Γ_t . Section 4.2 and Appendix A.3 allow Γ_t to be any price-based functional of the sampled marks. The proofs use only the local directional sensitivity of Γ_t to perturbations of the sampled prices.

2.4 Constrained traders

A continuum of traders $i \in [0, 1]$ face initial margin. Trader i has equity $E(i) > 0$ and must satisfy

$$|x_t(i)| M_t \leq E(i). \quad (5)$$

When M_{t+1} is posted, any trader who violates (5) must reduce exposure at $t + 1$ until the constraint holds; this adjustment is forced liquidation. For transparency, the main text uses a target-with-cap rule. Let $\bar{x}_t \in \mathbb{R}$ be a public target chosen from the pre-trade state S_t (Section 2.1), before the strategic trade and before P_t is realized. The realized position is the target clipped by the margin cap:

$$x_t(i) = \pi(\bar{x}_t, E(i)/M_t) = \text{sgn}(\bar{x}_t) \min \left\{ |\bar{x}_t|, \frac{E(i)}{M_t} \right\}. \quad (6)$$

Aggregate constrained exposure is

$$X_t = X(M_t, \bar{x}_t) = \int_0^1 \text{sgn}(\bar{x}_t) \min \left\{ |\bar{x}_t|, \frac{E(i)}{M_t} \right\} di. \quad (7)$$

When M_{t+1} rises, caps $E(i)/M_{t+1}$ tighten; if targets remain high, X_{t+1} falls and the reduction is liquidation order flow at $t + 1$. The analysis does not assume non-strategic constrained traders. It uses only one local object: the state-dependent slope of forced flow with respect to the next posted margin in states where the constraint binds. Define

$$c_X(S_t) = - \left. \frac{\partial X(M, \bar{x}_{t+1})}{\partial M} \right|_{M=M_{t+1}}, \quad (8)$$

where the derivative is evaluated at the benchmark path and holds fixed the constrained sector's period- $t + 1$ targets chosen from S_{t+1} before enforcement. Under the target-with-cap rule, $c_X(S_t) > 0$ whenever a positive mass of traders is margin-binding. Appendix B replaces the target rule by a forward-looking constrained sector that anticipates the update and chooses positions optimally. In that equilibrium the constrained block enters the mechanism only through the local liquidation slope

$$c_X^{\text{eff}}(S_t) = - \left. \frac{\partial X^*(M; S_t)}{\partial M} \right|_{M=M_{t+1}}.$$

Once positions are endogenized, the relevant object is not whether traders are strategic or anticipate the call, but whether the requirement is locally binding in the states of interest. This is exactly what the effective slope $c_X^{\text{eff}}(S_t)$ measures: it is the local elasticity of aggregate forced adjustment with respect to M_{t+1} after whatever anticipatory repositioning or within-window re-optimization is allowed. For design purposes, these are the states that matter: when the constraint binds, the rule mechanically induces forced adjustment and price pressure, so implementability must be assessed there rather than in slack states. Anticipation can attenuate this elasticity, but it does not change the logic of the theorem: the amplification condition and the implied slope-cap restriction are stated in terms of $c_X^{\text{eff}}(S_t)$ and therefore apply whenever $c_X^{\text{eff}}(S_t) > 0$. Under standard concavity, $c_X^{\text{eff}}(S_t) = 0$ only in states in which no participating trader is locally margin-binding at M_{t+1} , i.e., the requirement

is locally slack for the relevant set (or participation is negligible) in that state.⁶

2.5 Strategic trader and round trips

A strategic trader (or manipulator) chooses a predictable trade sequence a_0, \dots, a_T over a fixed horizon T , where $a_t > 0$ corresponds to a buy and $a_t < 0$ a sell. A round trip is a trade sequence with net-zero inventory at the end of the horizon:

$$\sum_{t=0}^T a_t = 0. \quad (9)$$

The resulting inventory process is $y_t = \sum_{s=0}^t a_s$. Let P_t denote the transaction price at time t . The trader's trading profit is

$$\Pi_T(a_{0:T}) = - \sum_{t=0}^T a_t P_t. \quad (10)$$

We allow an optional quadratic inventory penalty $\kappa \geq 0$. The resulting inventory-adjusted profit is

$$\Pi_T^\kappa(a_{0:T}) = \Pi_T(a_{0:T}) - \kappa \sum_{t=0}^{T-1} y_t^2. \quad (11)$$

Finally, we optionally impose a hard funding constraint: there exists $\bar{F} > 0$ such that the trader's margin outlay never exceeds \bar{F} :

$$M_t |y_t| \leq \bar{F} \quad \text{for all } t. \quad (12)$$

This constraint only restricts the maximum admissible trigger size and plays no role in the local gain mechanism.⁷

2.6 Information and equilibrium

A constrained trader i follows a Markov policy $\sigma_i^C : \mathcal{S} \rightarrow \mathbb{R}$ mapping the public state space $\mathcal{S} = \mathbb{R} \times \mathbb{R}^m \times (0, \infty)$ into an end-of-period position satisfying (5). Fix a benchmark policy profile $\{\sigma_i^C\}_{i \in [0,1]}$.

The strategic trader is introduced only as a deviator. Along the benchmark path he does not trade. Starting from a benchmark state S_t , we ask whether there exists an admissible finite-horizon round trip with strictly positive conditional expected profit generated by its effect on the next margin update, holding primitives, the constrained-trader policies, and the margin rule fixed. Since a single profitable deviation violates sequential optimality, exhibiting one such round trip is sufficient to reject manipulation-proofness at that state (Definition 3). Appendix B provides a microfoundation that endogenizes the constrained sector response through the equilibrium slope c_X^{eff} .

⁶Empirically, binding collateral and funding constraints are pervasive. See, among others, [Murphy et al. \(2014, 2016\)](#); [Aldasoro et al. \(2023\)](#); [King et al. \(2023\)](#) for evidence that margin calls generate large, state-dependent liquidity demands, and [Coval and Stafford \(2007\)](#); [Duarte and Eisenbach \(2013\)](#); [Greenwood et al. \(2015\)](#) for evidence on forced sales, fire-sale externalities, and price spillovers when constraints bind.

⁷Appendix A.5 adds proportional execution and linear funding wedges; these only add a minimum quantity feasibility restriction.

We write $(\cdot)^{(0)}$ for benchmark (no-deviation) objects. Fix S_t . Let $P_t^{(0)}$ be the period- t transaction price under the benchmark, and define the last return $R_t^{(0)} := P_t^{(0)} - P_{t-1}$. Let $RV_t^{(0)}$ be realized variance computed from the sampled marks $(P_{t-m}, \dots, P_{t-1}, P_t^{(0)})$, and let $M_{t+1}^{(0)} := g(RV_t^{(0)})$ be the posted margin. Because $\{\sigma_i^C\}$ is Markov in S_t and price formation is (2), the benchmark objects $(P_t^{(0)}, R_t^{(0)}, RV_t^{(0)}, M_{t+1}^{(0)})$ are measurable functions of S_t . A state is reachable if it occurs along the benchmark evolution induced by $\{\sigma_i^C\}$, g , and (2). All local assumptions are imposed on reachable states.

Definition 1 (Markov equilibrium under the margin rule). Fix g and the risk statistic (4). A Markov equilibrium is a pair $(\{\sigma_i^C\}_{i \in [0,1]}, P)$ such that, for all t : (i) constrained traders choose $x_t(i) = \sigma_i^C(S_t)$ and $X_t = \int_0^1 \sigma_i^C(S_t) di$; (ii) the transaction price satisfies (2) with $Q_t = \Delta X_t$ along the equilibrium; (iii) the margin designer posts $M_{t+1}^{(0)} = g(RV_t^{(0)})$ with $RV_t^{(0)}$ computed from realized transaction prices.

3 Properties and main results

This section defines the three properties that motivate the design discussion and proves the main result.

3.1 Three desirable properties

We state local properties around reachable benchmark states. Theorem 1 is a non-joint-feasibility result: if risk sensitivity and liquidity continuity hold, we construct one admissible finite-horizon round trip with strictly positive conditional expected profit, so round-trip manipulation-proofness fails. Liquidity continuity is a market-functioning desideratum; it is not the property contradicted in the proof.

Definition 2 (Risk sensitivity). A margin rule $M_{t+1} = g(\Gamma_t)$ is risk-sensitive on an interval $I \subset \mathbb{R}_+$ if g is differentiable on I and there exists $c_g > 0$ such that $g'(r) \geq c_g$ for all $r \in I$.

Risk sensitivity requires nontrivial local pass-through from the measured risk input into posted margin, as in risk-based margining and incentive-alignment arguments (Biais et al., 2016).

Definition 3 (Manipulation-proofness). A mechanism is round-trip manipulation-proof at a state S_t if for every admissible round trip $a_{t:t+T}$,

$$\mathbb{E}[\Pi_T^\kappa(a_{t:t+T}) \mid S_t] \leq 0.$$

This is the standard no-manipulation requirement in price-impact settings: profitable round trips correspond to manipulation/dynamic arbitrage and are ruled out by canonical no-manipulation conditions (Huberman and Stanzl, 2004).

Definition 4 (Liquidity continuity). Fix a Markov equilibrium (Definition 1). The mechanism has liquidity continuity if the mapping $S_t \mapsto (P_t^{(0)}, X_t^{(0)})$ is continuous on the set of reachable states and, for each reachable S_t , the map

$$\varepsilon \mapsto (P_t(\varepsilon), X_t(\varepsilon), P_{t+1}(\varepsilon), X_{t+1}(\varepsilon))$$

is continuous at $\varepsilon = 0$, where $P_s(\varepsilon)$ and $X_s(\varepsilon)$ are the outcomes when the strategic trader submits an additional trade ε at time t (relative to the benchmark with zero strategic trade) and all primitives are held fixed.

Liquidity continuity rules out endogenous liquidity cliffs: arbitrarily small within-window perturbations cannot trigger jumps in prices or aggregate positions. It is a continuity requirement, not a bound on steepness; discontinuities typically come from kinks, thresholds, or stepwise overlays in the composite mapping from sampled marks into posted margin (Brunnermeier and Pedersen, 2009; Shleifer and Vishny, 2011).

3.2 Assumptions

Fix a reachable benchmark state S_t and hold primitives and other agents' policies fixed. We construct a local two-step deviation. A trade at t perturbs the last sampled mark P_t , shifts the risk input (here RV_t), and moves the posted margin M_{t+1} . If the requirement binds, the higher M_{t+1} forces liquidation at $t + 1$, which moves P_{t+1} through price impact; reversing the initial trade can monetize this predictable liquidation-driven price component. The proof bounds each local pass-through. Appendix A states the same conditions for a generic price-based input and non-smooth overlays using one-sided slopes and directional derivatives. Multi-period trigger-and-reverse attacks against fixed deterministic rulebooks are characterized in Nicolai (2026); here we identify the local design trade-off that makes such vulnerabilities unavoidable under risk sensitivity and continuity near binding states.

Assumption 1 (Local slope of the margin rule). There exists an interval $I = [\underline{r}, \bar{r}] \subset \mathbb{R}_+$ such that g is differentiable on I and

$$g'(r) \in [c_g, C_g] \quad \text{for all } r \in I,$$

with $0 < c_g \leq C_g < \infty$.

Assumption 2 (Local sensitivity of the risk input). At S_t , $RV_t^{(0)} \in \text{int}(I)$ and $|R_t^{(0)}| \geq \underline{R}$ for some $\underline{R} > 0$. The strategic trader can trade in either direction with capacity $\bar{a} > 0$. In the explicit construction we take $R_t^{(0)} < 0$; the case $R_t^{(0)} > 0$ is obtained by reversing signs.

Assumption 3 (Local liquidation slope under binding margin). At S_t with $RV_t^{(0)} \in \text{int}(I)$, there exist $c_X > 0$ and $\delta_X > 0$ such that, holding the period- t target \bar{x}_t fixed, the mapping $M \mapsto X(M, \bar{x}_t)$ is continuously differentiable on $[M_t, M_t + \delta_X]$ and satisfies

$$-\frac{\partial X}{\partial M}(M, \bar{x}_t) \geq c_X \quad \text{for all } M \in [M_t, M_t + \delta_X].$$

Assumption 3 is the local statement that when a positive mass of traders is margin-binding, a higher posted margin forces a first-order reduction in aggregate exposure. Under the target-with-cap policy it holds in any such binding state. Appendix B endogenizes targets in a forward-looking equilibrium; in that case the same condition holds with c_X replaced by the equilibrium slope $c_X^{\text{eff}}(S_t)$, the derivative of equilibrium aggregate exposure with respect to the next posted margin.

Assumption 4 (Price impact). $\alpha > 0$.

3.3 Auxiliary lemmas

We isolate four links used in the two-step deviation: (i) how RV_t moves when the last sampled mark P_t moves, (ii) how g maps this into M_{t+1} , (iii) how a higher margin forces liquidation when constraints bind, and (iv) how that forced flow moves P_{t+1} under linear impact.

Lemma 1 (Sensitivity of realized variance to the last mark). *Fix $(P_{t-m}, \dots, P_{t-1})$ and a benchmark last price $P_t^{(0)}$. Let $R_t^{(0)} = P_t^{(0)} - P_{t-1}$. The map $p \mapsto RV_t(P_{t-m}, \dots, P_{t-1}, p)$ in (4) is differentiable and*

$$\left. \frac{\partial RV_t}{\partial P_t} \right|_{P_t=P_t^{(0)}} = 2R_t^{(0)}.$$

If $|R_t^{(0)}| \geq \underline{R} > 0$, then for any ΔP with $\text{sgn}(\Delta P) = \text{sgn}(R_t^{(0)})$,

$$RV_t(P_t^{(0)} + \Delta P) - RV_t(P_t^{(0)}) \geq 2\underline{R}|\Delta P|.$$

Proof. Holding $(P_{t-m}, \dots, P_{t-1})$ fixed, only $(P_t - P_{t-1})^2$ depends on P_t , so

$$\frac{\partial RV_t}{\partial P_t} = 2(P_t - P_{t-1}),$$

and evaluating at $P_t^{(0)}$ gives $2R_t^{(0)}$. Also,

$$RV_t(P_t^{(0)} + \Delta P) - RV_t(P_t^{(0)}) = (R_t^{(0)} + \Delta P)^2 - (R_t^{(0)})^2 = 2R_t^{(0)}\Delta P + (\Delta P)^2.$$

If $\text{sgn}(\Delta P) = \text{sgn}(R_t^{(0)})$, then $R_t^{(0)}\Delta P = |R_t^{(0)}||\Delta P|$ and $(\Delta P)^2 \geq 0$, so the difference is at least $2|R_t^{(0)}||\Delta P| \geq 2\underline{R}|\Delta P|$. \square

Lemma 2 (Pass-through from RV_t to margin). *Under Assumption 1, for any $RV, RV' \in I$ with $RV' \geq RV$,*

$$g(RV') - g(RV) \geq c_g (RV' - RV).$$

Proof. By the mean value theorem, $g(RV') - g(RV) = g'(\xi)(RV' - RV)$ for some ξ between RV and RV' . Assumption 1 gives $g'(\xi) \geq c_g$. \square

Lemma 3 (Liquidation induced by a margin increase). *Under Assumption 3, there exists $\delta_M > 0$ such that for all $\Delta M \in (0, \delta_M)$,*

$$L(\Delta M) = X(M_t, \bar{x}_t) - X(M_t + \Delta M, \bar{x}_t) \geq c_X \Delta M,$$

and $\Delta M \mapsto L(\Delta M)$ is continuous on $(0, \delta_M)$.

Proof. Write $X(M) = X(M, \bar{x}_t)$ and set $\delta_M = \delta_X$. For $\Delta M \in (0, \delta_M)$, the mean value theorem gives $X(M_t + \Delta M) - X(M_t) = X'(\xi)\Delta M$ for some $\xi \in (M_t, M_t + \Delta M)$. Assumption 3 implies $X'(\xi) \leq -c_X$, so $L(\Delta M) \geq c_X \Delta M$. Continuity follows from continuity of X on $[M_t, M_t + \delta_M]$. \square

Lemma 4 (Liquidation-induced price change under linear impact). *Fix $t+1$. Suppose the margin update induces additional forced liquidation $L > 0$ at $t+1$ relative to the benchmark. Let $\Delta X_{t+1}^{(0)}$ be benchmark*

constrained order flow at $t + 1$, so realized constrained order flow is $\Delta X_{t+1} = \Delta X_{t+1}^{(0)} - L$. If the strategic trader buys $q \in [0, \bar{a}]$ at $t + 1$, then

$$P_{t+1} = V_{t+1} + \alpha(\Delta X_{t+1}^{(0)} - L + q).$$

Holding $\Delta X_{t+1}^{(0)}$ fixed, the liquidation component is $-\alpha L$.

Proof. Substitute $Q_{t+1} = \Delta X_{t+1} + q = \Delta X_{t+1}^{(0)} - L + q$ into (2). \square

3.4 An explicit round trip

This section provides a transparent two-period deviation that exposes the feedback loop. Fix a time t and consider a trader who sells $q > 0$ at t and buys q at $t + 1$. The trade sequence is $(a_t, a_{t+1}) = (-q, q)$, and the end-of-horizon inventory is zero. For the realized-variance specification (4), a sale at t changes the time- t return from $R_t^{(0)}$ to $R_t = R_t^{(0)} - \alpha q$ and hence changes realized variance by

$$\Delta RV_t = RV_t - RV_t^{(0)} = 2R_t^{(0)}(-\alpha q) + (\alpha q)^2. \quad (13)$$

When $R_t^{(0)} < 0$, a sale increases realized variance, which increases margin when g is risk-sensitive. To simplify notation, define the local sensitivity constant

$$c_{RV} = 2|R_t^{(0)}|.$$

Then, under the sign choice $R_t^{(0)} < 0$, we have $\Delta RV_t \geq c_{RV}\alpha q$. Combining this with risk sensitivity of g and the liquidation elasticity yields a lower bound on forced liquidation at $t + 1$ of the form

$$L \geq c_X c_g c_{RV} \alpha q.$$

Define the corresponding feedback gain

$$G_t = \alpha c_X c_g c_{RV}. \quad (14)$$

The next lemma computes the inventory-adjusted profit from the two-period round trip.

Lemma 5 (Profit from the two-period round trip). *Fix a reachable state S_t satisfying Assumptions 1–4. Consider the two-period round trip $(a_t, a_{t+1}) = (-q, q)$ with $q > 0$. Suppose q is admissible, meaning that all locality restrictions hold (including $RV_t + \Delta RV_t \in (\underline{r}, \bar{r})$ and $M_{t+1} + \Delta M_{t+1} \in I$) and that the hard funding constraint (12) holds if imposed. Then the expected inventory-adjusted profit satisfies*

$$\mathbb{E}[\Pi_1^\kappa \mid S_t] \geq \alpha q (G_t q - 2q) - \kappa q^2 = (\alpha(G_t - 2) - \kappa) q^2. \quad (15)$$

In particular, if $\alpha(G_t - 2) > \kappa$, then $\mathbb{E}[\Pi_1^\kappa \mid S_t] > 0$ for every admissible $q > 0$ (equivalently, for every $q \in (0, q_{\max}]$ in the local range).

Proof. Under the round trip, $\Pi_1^\kappa = -a_t P_t - a_{t+1} P_{t+1} - \kappa |y_t|^2$. Since $a_t = -q$, $a_{t+1} = q$, and $y_t = -q$,

this becomes

$$\Pi_1^\kappa = q(P_t - P_{t+1}) - \kappa q^2.$$

Write prices under the deviation relative to the benchmark: $P_t = P_t^{(0)} - \alpha q$ and $P_{t+1} = P_{t+1}^{(0)} + \alpha(q - L)$. Hence

$$P_t - P_{t+1} = (P_t^{(0)} - P_{t+1}^{(0)}) + \alpha(L - 2q).$$

Assume the benchmark has no predictable drift from t to $t + 1$, i.e., $\mathbb{E}[P_{t+1}^{(0)} | S_t] = P_t^{(0)}$, so $\mathbb{E}[P_t^{(0)} - P_{t+1}^{(0)} | S_t] = 0$. Taking conditional expectations gives

$$\mathbb{E}[\Pi_1^\kappa | S_t] = \alpha q(L - 2q) - \kappa q^2.$$

Substituting the lower bound $L \geq c_X c_g c_{RV} \alpha q$ yields (15). \square

We emphasize that the deviation is local: admissibility requires that the trade be small enough to keep (RV_t, M_{t+1}) within the risk-sensitive region and within the model's local elasticity bounds. In the present special case, this restriction is $2 \left| R_t^{(0)} \right| \alpha q + \alpha^2 q^2 \leq \bar{r} - RV_t^{(0)}$ when the deviation raises RV_t . Lemma 5 shows that when the local gain condition holds, the expected profit is strictly positive for arbitrarily small admissible trigger sizes.

3.5 Main Theorem

Lemma 5 gives a transparent two-period deviation and isolates a local gain condition under which the expected profit is strictly positive for arbitrarily small triggers. Call a trigger size $q > 0$ admissible at S_t if all locality restrictions required by Assumptions 1–4 (and, when imposed, the capacity and hard-funding constraints) continue to hold under the deviation. Let $q_{\max}(S_t) \in (0, \infty]$ denote the supremum of admissible trigger sizes at S_t , so that every $q \in (0, q_{\max}(S_t)]$ is admissible.⁸

Theorem 1 (Impossibility). *Suppose Assumptions 1–4 hold at a reachable state S_t (and the benchmark drift condition used in Lemma 5 holds). If $\alpha(G_t - 2) > \kappa$ and $q_{\max}(S_t) > 0$, then the mechanism is not round-trip manipulation-proof at S_t : for every admissible $q \in (0, q_{\max}(S_t)]$, the two-period round trip has strictly positive conditional expected inventory-adjusted profit. Consequently, risk sensitivity (Definition 2), round-trip manipulation-proofness (Definition 3), and liquidity continuity (Definition 4) cannot hold simultaneously at S_t .*

Proof. Fix any admissible $q \in (0, q_{\max}(S_t)]$. By Lemma 5,

$$\mathbb{E}[\Pi_1^\kappa | S_t] \geq (\alpha(G_t - 2) - \kappa) q^2.$$

Under $\alpha(G_t - 2) > \kappa$, the right-hand side is strictly positive for every $q > 0$, so $\mathbb{E}[\Pi_1^\kappa | S_t] > 0$. This violates round-trip manipulation-proofness (Definition 3), hence the three properties cannot all hold at S_t . \square

⁸A sufficient admissible range can be constructed explicitly by bounding how far the deviation can move the risk input and next-period requirement; one convenient choice is $q_{\max} = \min\{q_I, q_M, \bar{a}, \bar{F}/M_t\}$ with q_I ensuring $RV_t + \Delta RV_t \leq \bar{r}$ and q_M ensuring $M_{t+1}^{(0)} + \Delta M_{t+1} \leq M_{t+1}^{(0)} + \delta_X$.

Theorem 1 is a local incentive-compatibility result for binding, price-based margin rules. In a benchmark with no predictable fundamental drift, sufficiently small round trips should lose to price impact. The theorem shows why this benchmark breaks once transaction-based marks enter the rule-book for next-period feasibility. A trade inside the sampling window moves a sampled transaction price, shifts the measured risk input Γ_t , and therefore shifts the posted requirement $M_{t+1} = g(\Gamma_t)$ at first order. When the requirement binds for a nontrivial constrained sector, this marginal tightening induces mechanical deleveraging, which enters price formation through impact. The deviator completes the round trip by providing liquidity against that predictable, rule-induced order flow. Profitability therefore comes from trading against a scheduled reset of a binding constraint, not from forecasting fundamentals or sustaining a mispricing.

Two clarifications address common misreadings. First, the deviation does not rely on large trades or large price moves: under the gain condition, expected inventory-adjusted profits are strictly positive for every admissible trigger size $q \in (0, q_{\max}]$. Shrinking q shrinks both the deviator's direct impact costs and the induced liquidation in the same local order, leaving the sign of expected profit unchanged. Second, the deviation does not rely on making a slack constraint bind. The theorem is evaluated at a reachable state in which some participants are already locally constrained, so a marginal increase in M_{t+1} induces a marginal reduction in aggregate exposure without any discrete regime switch.

The proof is written as a two-period round trip because manipulation-proofness is an incentive constraint: a single profitable deviation is sufficient to violate it at the evaluated state. In applications, requirements reset repeatedly, so the same economics extend across update cycles. Nicolai (2026) provides a dynamic extension: the optimal finite-horizon trigger-and-reverse strategy decomposes into (i) intermediation of the predictable post-update forced flow and (ii) trades within the sampling window that shift the measured statistic and reshape the magnitude and timing of subsequent forced rebalancing. The relevant primitives are the local sensitivity of the statistic to sampled prints and the local slope of g in binding regions; with convex inputs such as volatility-based measures, this within-window component can be implemented with alternating-sign trades that keep net inventory small while increasing the statistic (Nicolai, 2026).

Interpreted as a design restriction, Theorem 1 pins down what a binding, price-based margin rule can and cannot do locally. In the region where the rule binds, local risk sensitivity means that small changes in sampled transaction prices pass through at first order into next-period requirements. If the resulting amplification chain from marks to requirements to forced flow to prices is strong enough, then no rule can simultaneously be (i) locally risk-sensitive, (ii) liquidity-continuous, and (iii) round-trip manipulation-proof. Put differently, manipulation-proofness in binding states requires limiting short-horizon pass-through from transaction-based inputs into posted requirements. This gives a structural rationale for practical devices that attenuate local pass-through, such as slope limits, smoothing, buffers, and pooling regions.

Corollary 1 (Implied local slope cap). *Fix a reachable binding state S_t satisfying Assumptions 2–4 and suppose g is differentiable at the benchmark input $RV_t^{(0)}$. If liquidity continuity holds at S_t and the mechanism*

is round-trip manipulation-proof at S_t , then the local pass-through slope must satisfy

$$g'(RV_t^{(0)}) \leq \frac{2\alpha + \kappa}{\alpha^2 c_X c_{RV}} = \frac{2\alpha + \kappa}{2\alpha^2 c_X |R_t^{(0)}|}, \quad (16)$$

where $c_{RV} = 2 |R_t^{(0)}|$ is the local sensitivity constant of the risk-measure. In particular, since $|R_t^{(0)}| \leq \sqrt{RV_t^{(0)}}$, a sufficient implementable envelope stated as a function of the variance-type input alone is

$$g'(r) \leq \frac{2\alpha + \kappa}{2\alpha^2 c_X \sqrt{r}}. \quad (17)$$

Equivalently, writing $\sigma = \sqrt{r}$ and $\tilde{g}(\sigma) = g(\sigma^2)$, the envelope implies the constant volatility-scale cap

$$\tilde{g}'(\sigma) \leq \frac{2\alpha + \kappa}{\alpha^2 c_X}. \quad (18)$$

Proof. Suppose, toward a contradiction, that $g'(RV_t^{(0)}) > (2\alpha + \kappa)/(\alpha^2 c_X c_{RV})$. By continuity of g' , there exist an interval I containing $RV_t^{(0)}$ and a constant $\underline{c}_g > 0$ such that $g'(r) \geq \underline{c}_g$ for all $r \in I$ and $\underline{c}_g > (2\alpha + \kappa)/(\alpha^2 c_X c_{RV})$. Choose $q > 0$ small enough that the deviation keeps the perturbed input inside I and within the local elasticity region of Assumption 3. Then Lemma 5 implies

$$\mathbb{E}[\Pi_1^\kappa | S_t] \geq (\alpha(\alpha c_X \underline{c}_g c_{RV} - 2) - \kappa) q^2 > 0,$$

which contradicts round-trip manipulation-proofness at S_t . Hence (16) must hold. The envelope (17) follows from $|R_t^{(0)}| \leq \sqrt{RV_t^{(0)}}$. Finally, for $\tilde{g}(\sigma) = g(\sigma^2)$ we have $\tilde{g}'(\sigma) = 2\sigma g'(\sigma^2)$, and (18) follows from (17). \square

Section 3.6 translates the local gain condition of Theorem 1 and the implied slope-cap logic of Corollary 1 into observable objects and conservative magnitude screens, keeping the focus on what can be assessed without position-level data.

3.6 Back-of-the-envelope magnitudes for the feedback gain

The magnitude implication in Theorem 1 can be summarized by two objects that map directly into data. First, the forced-liquidation multiple $G = L/q$, where q is the trigger trade at the sampling time t and L is the additional forced liquidation at $t + 1$ induced by the higher posted requirement. Under linear impact, G is the number of units liquidated per unit triggered. Lemma 5 implies that the two-period trigger-and-reverse deviation is profitable (in inventory-adjusted terms) when $\alpha(G - 2) > \kappa$.

Second, the sampled-mark distortion. Let δ_{bps} denote the distortion of the sampled mark created by the trigger trade, in bps of notional. With $P_t = P_t^{(0)} - \alpha q$, the relative price displacement is $\alpha q / P_t$, so

$$\delta_{\text{bps}} = 100 \delta_{\%} \quad \text{where} \quad \delta_{\%} = 100 \frac{\alpha q}{P_t}.$$

A conservative way to discipline δ_{bps} is to express the trigger size in participation units $f = (qP_t)/ADV$,

where ADV is average daily traded value. Using the standard square-root scaling of impact in participation units, we use the benchmark mapping

$$\delta_{\text{bps}}(f) \approx 200\sqrt{f},$$

with f measured as a fraction of full-day ADV .⁹ Using full-day ADV is conservative for margining because marks are typically computed from a much shorter sampling window, so effective participation within the window is higher than f . Lemma 5 implies the profit-per-notional lower bound in percent units,

$$\pi_{\%} = 100 \frac{\mathbb{E}[\Pi_1 | S_t]}{qP_t} \geq (G - 2) \delta_{\%},$$

equivalently, in bps units,

$$\pi_{\text{bps}} = 100 \pi_{\%} \geq (G - 2) \delta_{\text{bps}}.$$

Normalizing by peak margin usage mqP_t , where $m = M_t/P_t$ is the margin ratio, the implied return-on-margin is

$$ROI_{\%} = 100 \frac{\mathbb{E}[\Pi_1 | S_t]}{m q P_t} \geq \frac{(G - 2) \delta_{\text{bps}}}{100 m}.$$

For example, if $f = 1\%$ so that $\delta_{\text{bps}} \approx 20$, and $G = 6$, then $\pi_{\text{bps}} \geq 80$ bps in one update and $ROI_{\%} \geq 80/(100m)$; for $m = 7\%$, this is about 11% return on posted margin for a single reset.

A simple accounting identity rewrites $G = L/q$ in terms of observable quantities. Let X denote the constraint-sensitive sector's position and define the position-to-margin elasticity $\varepsilon_{X,M} = -d \log X / d \log M$. For a relative requirement change $\Delta M/M$,

$$\frac{L}{X} \approx \varepsilon_{X,M} \frac{\Delta M}{M}.$$

Define the turnover ratio $\tau = (XP_t)/ADV$. Since $f = (qP_t)/ADV$, we have $q/X = f/\tau$, and therefore

$$G = \frac{L}{q} \approx \varepsilon_{X,M} \frac{\Delta M}{M} \frac{\tau}{f}.$$

This highlights why the theorem has bite in binding states: (i) short-horizon margin moves $\Delta M/M$ can be large under standard risk-based models; (ii) impact and thus δ_{bps} are larger in stress; (iii) G grows when the constrained sector is large relative to market depth (τ high) and when the trigger is a small fraction of volume (f small). Appendix D provides the empirical calibration of $\varepsilon_{X,M}$, $\Delta M/M$, and $\delta_{\text{bps}}(f)$.

Table 1 interprets the two-period trigger-and-reverse deviation in economically familiar units. Panel A reports the forced-liquidation multiple $G = L/q$: the number of units of constrained-sector

⁹The square-root form is a standard empirical benchmark for how implementation shortfall scales with participation (Frazzini et al., 2018; Kyle and Obizhaeva, 2016). To fix a conservative scale, start from an empirical one-way average cost curve $\text{cost}_{\text{bps}}(f) \approx k\sqrt{f}$, where f is the trade size as a fraction of daily volume. Estimates in Frazzini et al. (2018) suggest k on the order of 100 bps, so a 1% of ADV trade has average one-way cost about $100\sqrt{0.01} \approx 10$ bps. The distortion that matters for a sampled transaction mark is closer to an instantaneous price displacement at the sampling time. Under the standard constant-rate execution with linear instantaneous impact, impact ramps from 0 to a terminal (peak) displacement, so the time-average displacement paid is half the peak; equivalently, peak displacement is about twice the average cost. We therefore use the conservative mapping $\delta_{\text{bps}}(f) \approx 2k\sqrt{f} \approx 200\sqrt{f}$.

Table 1 Illustrative profitability

The table reports illustrative one-update profitability for the trigger-and-reverse deviation under conservative benchmark parameters: $\epsilon_{X,M} = 0.225$ (Hedegaard, 2014), $\tau = 1$, and $\delta_{\text{bps}}(f) \approx 200\sqrt{f}$ (Frazzini et al., 2018). Columns correspond to margin jumps from $m_0 = 5\%$ to $m_1 \in \{6\%, 7\%, 8\%\}$ (so $\Delta M/M \in \{20\%, 40\%, 60\%\}$). Panel A reports the implied forced-liquidation multiple $G = L/q$. Panel B reports the corresponding lower bounds for profit in bps of trigger notional, π_{bps} , and return on posted margin, $ROI\%$, computed from $\pi_{\text{bps}} \geq (G - 2)\delta_{\text{bps}}$ and $ROI\% \geq \pi_{\text{bps}}/(100 m_1)$.

	$m_0 = 5\% \rightarrow m_1 = 6\%$	$m_0 = 5\% \rightarrow m_1 = 7\%$	$m_0 = 5\% \rightarrow m_1 = 8\%$
Panel A: liquidation multiple $G = L/q$			
$f = 0.5\%$	9.0	18.0	27.0
$f = 1\%$	4.5	9.0	13.5
$f = 2\%$	2.25	4.5	6.75
Panel B: profitability lower bounds			
	$\pi_{\text{bps}} \geq (G - 2)\delta_{\text{bps}} \quad \text{and} \quad ROI\% \geq \pi_{\text{bps}}/(100 m_1)$		
	$\pi_{\text{bps}}/ROI\%$	$\pi_{\text{bps}}/ROI\%$	$\pi_{\text{bps}}/ROI\%$
$f = 0.5\%$	99.0/16.5	226.3/32.3	353.6/44.2
$f = 1\%$	50.0/8.3	140.0/20.0	230.0/28.8
$f = 2\%$	7.1/1.2	70.7/10.1	134.4/16.8

liquidation induced at $t + 1$ per unit of trigger trade executed at the sampling time t . Values well above 2 mean that the forced-flow response overwhelms the trader’s own round-trip impact loss, which is the profitability screen in Lemma 5. Panel B translates the same cases into profitability. Each cell reports a lower bound on (i) profit in bps of trigger notional, π_{bps} , and (ii) return on posted margin, $ROI\%$, where the denominator is the post-update margin $m_1 q P_t$. For example, with $f = 1\%$ participation and a margin jump from 5% to 7%, the table implies $G = 9$ and a one-update profit bound of at least 140 bps on the trigger notional, corresponding to at least 20% return on posted margin in a single update. Even in the mildest scenario ($5\% \rightarrow 6\%$) at $f = 1\%$, the bound is 50 bps and 8.3% on margin.

These numbers are deliberately conservative in three ways. First, $\delta_{\text{bps}}(f)$ is disciplined using full-day *ADV* rather than the typically thinner settlement or snapshot window. Second, the ROI denominator uses the higher post-update margin m_1 rather than the pre-update margin m_0 . Third, the table reports only the direct one-step margin-update channel captured by Lemma 5. It therefore understates profits from dynamic attacks that repeat the mechanism across update times, exploit state dependence of impact and margin sensitivity in stress, or combine trigger and harvest legs across contracts under portfolio margining (Section 5). The calibration inputs behind Table 1 (elasticity $\epsilon_{X,M}$, stress-time margin moves $\Delta M/M$, and the impact mapping $\delta_{\text{bps}}(f)$) and the full sensitivity grids over $(\epsilon_{X,M}, \tau, f)$ are collected in Appendix D.

3.6.1 How large are one-update margin moves in practice?

In Table 1 we treat relative increases in posted margin of 20%, 40%, and 60% as benchmark scenarios. Risk-based initial margin models can generate short-horizon margin calls of this order even when anti-procyclicality tools are used, because APC tools mitigate procyclicality mainly by raising mar-

gins in tranquil times (buffers and floors). This raises the level of posted collateral and mechanically reduces ROI, but it does not eliminate large stress-time margin moves unless floors and buffers are made so large that risk sensitivity is materially reduced.

Simulations in [Murphy et al. \(2014\)](#) report large short-horizon margin calls as a fraction of portfolio value. Reading these as changes in the required margin ratio, a representative starting margin around 5% ([Hedegaard, 2014](#)) implies relative increases on the order of 20% over one day and around 40% over multi-day horizons under standard specifications; if the starting point is an unusually tranquil period with laxer posted margins, the same absolute call translates into even larger relative jumps. Stress-episode evidence makes the same point. In a detailed case study of S&P 500 futures around March 2020, [Gurrola-Perez \(2020\)](#) documents very large short-horizon increases under filtered historical simulation VaR and shows that materially reducing the worst multi-day increases requires large APC floors, at the cost of substantially higher margins in normal times. The design implication matches the theorem’s tension: limiting stress-time jumps requires slope-limiting overlays that attenuate local pass-through, moving the rule away from local risk sensitivity.

The same logic extends beyond CCPs. Any price-based rule that maps sampled marks into a binding requirement can exhibit large short-horizon adjustments in binding states, and those are precisely the states in which the local profitability screen in [Lemma 5](#) has economic bite.

4 Extensions

For exposition, the main text uses a deliberately bare-bones specification: linear price impact, a realized-variance risk input, and a simple margin-cap demand rule. Here we record extensions useful for interpretation.

4.1 Continuous trading

The discrete-time indexing in [Sections 2–3](#) indexes margin update times. Trading can be continuous between updates without changing the argument, because the risk input is computed from a sampling grid of marks and applies the margin update at discrete times. Fix an update interval $h > 0$ and update times $t_n = nh$, $n = 0, 1, 2, \dots$. Let $P_n = P_{t_n}$ and $V_n = V_{t_n}$. Let ΔX_n denote the constrained traders’ net order flow aggregated over $(t_{n-1}, t_n]$, and let a_n denote the strategic trader’s net order flow over the same interval. At update times, transaction prices satisfy

$$P_n = V_n + \alpha(\Delta X_n + a_n), \quad (19)$$

realized variance is computed from the last m sampled marks,

$$RV_n = \sum_{j=0}^{m-1} (P_{n-j} - P_{n-j-1})^2, \quad (20)$$

and next-period margin is

$$M_{n+1} = g(RV_n). \quad (21)$$

Constrained traders face the same margin constraint at update times and adjust mechanically when M_{n+1} is posted.

Proposition 1 (Continuous trading with discrete margin updates). Fix $h > 0$ and the update-time model (19)–(21). Suppose the local assumptions of Section 3.2 hold at some reachable update-time state, interpreted with (P_t, V_t, RV_t, M_t) replaced by (P_n, V_n, RV_n, M_n) . Suppose also that the local conditions of Lemma 5 hold at that state (with t replaced by n). Then the same two-step round trip (net selling over $(t_{n-1}, t_n]$ and net buying over $(t_n, t_{n+1}]$) delivers the same profit lower bound (15). In particular, whenever the amplification condition in Lemma 5 holds, the impossibility result of Theorem 1 applies at that update-time state.

Proof. Define the update-time state $S_n = (V_n, P_{n-m}, \dots, P_{n-1}, M_n)$. The price rule (19), the risk statistic (20), and the margin rule (21) have the same functional form as (2), (4), and (3) after replacing t by n . The argument in Lemma 5 is algebraic and uses only these objects together with the local demand and impact slopes. Applying the same steps at update times yields (15), and Theorem 1 follows in the same way. \square

Risk measures are typically computed from settlement prices, end-of-day marks, or specified intra-day snapshots. Even in a continuously traded market, the risk input is built from a discrete sample, and the margin update is a discrete event.¹⁰

4.2 Alternative risk inputs

We use realized variance in the main text because its local sensitivity to the last sampled mark is transparent. The mechanism does not rely on squaring returns. What matters is that the CCP computes a price-based risk input from sampled marks, and that some implementable perturbation of a sampled mark moves this input at first order.

4.2.1 A generic price-based risk functional

Let the CCP compute a scalar risk statistic from the sampled price vector,

$$\Gamma_t = \Gamma(P_{t-m}, \dots, P_t), \quad M_{t+1} = g(\Gamma_t).$$

The Appendix allows non-smooth Γ and non-smooth overlays inside g . Here we isolate the single local property that replaces the realized-variance sensitivity used in the main text.

¹⁰This extension is meant to address the continuous-trading concern directly: allowing continuous trading between update times does not weaken the result. For any fixed update interval $h > 0$, Proposition 1 applies verbatim, because the CCP still computes the risk input from a discrete set of marks and applies the margin update at discrete times. The only boundary case in which first-order leverage can disappear is an idealized diffusion limit that simultaneously (i) lets the sampling grid become arbitrarily fine, (ii) assumes continuous price paths, and (iii) holds the trader's per-update trade size bounded, so that $\Delta RV_n = 2R_n \Delta P_n + (\Delta P_n)^2$ with $R_n = O(\sqrt{h})$ and the linear term vanishes as $h \downarrow 0$. But that is not a continuous trading limit in the institutional sense; it requires a smooth rule to be recomputed and enforced essentially continuously from a continuously sampled path, with no discrete marks (settlements, auctions, snapshots) and no operational kinks (thresholds, rounding, stress overlays). Actual margin engines retain discrete marks and discrete enforcement even in large, liquid markets, so the empirically relevant case is precisely Proposition 1.

Assumption 5 (Local one-sided sensitivity of a generic risk statistic). There exists a reachable state S_t such that $\Gamma_t \in \text{int}(I_\Gamma)$ for some interval $I_\Gamma \subset \mathbb{R}_+$, and there exist a sign choice and a constant $c_\Gamma > 0$ such that, holding $(P_{t-m}, \dots, P_{t-1})$ fixed,

$$\Gamma(P_{t-m}, \dots, P_{t-1}, P_t + \Delta P) - \Gamma(P_{t-m}, \dots, P_{t-1}, P_t) \geq c_\Gamma |\Delta P|$$

for all sufficiently small ΔP of that sign.

Assumption 5 says that, at some reachable state, there is an implementable one-sided perturbation of the last sampled price that raises the CCP risk input at a nontrivial linear rate.

Theorem 2 (Impossibility for a locally sensitive price-based risk input). Replace (4) by $\Gamma_t = \Gamma(P_{t-m}, \dots, P_t)$ and (3) by $M_{t+1} = g(\Gamma_t)$. Replace Assumption 2 by Assumption 5. Assume g is risk-sensitive on an interval I_Γ containing Γ_t in its interior, and retain Assumptions 3–4 and the local conditions used in Lemma 5. Then, for the same two-period round trip (sell q at t and buy q at $t+1$) and for q small enough that the perturbed Γ_t remains in I_Γ ,

$$\mathbb{E}[\Pi_1^\kappa \mid S_t] \geq \alpha q (\alpha c_X c_g c_\Gamma q - 2q) - \kappa q^2 = (\alpha(\alpha c_X c_g c_\Gamma - 2) - \kappa) q^2.$$

In particular, if $\alpha c_X c_g c_\Gamma > 2 + \kappa/\alpha$, then $\mathbb{E}[\Pi_1^\kappa \mid S_t] > 0$ for every admissible q in the local range.

Proof. With $\Delta X_t = 0$, selling q at t moves the time- t transaction price by $\Delta P_t = -\alpha q$. By Assumption 5 (choosing the sign appropriately),

$$\Delta \Gamma_t \geq c_\Gamma |\Delta P_t| = c_\Gamma \alpha q.$$

Risk sensitivity of g on I_Γ yields $\Delta M_{t+1} \geq c_g \Delta \Gamma_t \geq c_g c_\Gamma \alpha q$, and Assumption 3 yields $L \geq c_X \Delta M_{t+1} \geq c_X c_g c_\Gamma \alpha q$. The profit identity in Lemma 5 gives $\mathbb{E}[\Pi_1^\kappa \mid S_t] = \alpha q (L - 2q) - \kappa q^2$, so substituting the lower bound for L yields the claim. \square

4.2.2 VaR/ES built from volatility

A common implementation sets margin proportional to a VaR or ES multiple of a volatility estimate, for example $\Gamma_t = z_p \hat{\sigma}_t$ or $\Gamma_t = k_p \hat{\sigma}_t$. If $\hat{\sigma}_t$ loads on recent squared returns with positive weight, then $\hat{\sigma}_t$ inherits a local first-order sensitivity to a perturbation of a sampled price whenever the relevant sampled return is not exactly zero. For the special case $\hat{\sigma}_t = \sqrt{RV_t/(m-1)}$,

$$\frac{\partial \hat{\sigma}_t}{\partial P_t} = \frac{1}{2\sqrt{(m-1)RV_t}} \cdot \frac{\partial RV_t}{\partial P_t} = \frac{R_t}{\sqrt{(m-1)RV_t}}.$$

This expression highlights a sharp state dependence: holding the most recent return R_t fixed, the local slope scales like $1/\sqrt{RV_t}$. Hence the same small perturbation of the last sampled mark has a much larger effect on $\hat{\sigma}_t$ after a quiet spell (low RV_t) than in an already volatile state. Put differently, low past measured risk is precisely when volatility-based risk inputs are locally steep: a single nontrivial return arriving after low realized variance mechanically produces a large change in $\hat{\sigma}_t$, and therefore in Γ_t and M_{t+1} .

Existing leverage-cycle and endogenous-risk work focuses on an equilibrium channel: when measured risk is low, constraints and funding terms loosen, balance sheets expand, and fragility accumulates (Adrian and Shin, 2010; Brunnermeier and Pedersen, 2009; Geanakoplos, 2010; Brunnermeier and Sannikov, 2014; Danielsson et al., 2004). Our point is different. We isolate a purely mechanical, local object for price-based rulebooks: the directional sensitivity of the posted requirement to a manipulable sampled mark. For volatility-based inputs this sensitivity is steepest precisely in tranquil states (low recent RV_t), so the states that look safest under the rulebook are also the states that are easiest to move at first order. To the best of our knowledge, this link between low measured risk and maximal first-order manipulability of price-based constraints is not made explicit in the existing leverage-cycle and procyclicality literatures. It is central for large, liquid markets, which spend long stretches in low-volatility regimes.

4.2.3 Scenario maxima, quantiles, and kinks

Many CCP engines compute a scalar risk number as a maximum over scenarios or as a high quantile of scenario losses:

$$\Gamma_t = \max_{s \in \mathcal{S}} l_s(P_{t-m:t}), \quad \text{or} \quad \Gamma_t = \text{Quantile}_p(l_1(P_{t-m:t}), \dots, l_N(P_{t-m:t})).$$

These mappings are typically non-differentiable when the identity of the worst scenario (or the binding order statistic) changes. Away from kink points, a single active scenario governs the local slope, and the analysis reduces to Theorem 2 with c_Γ equal to the one-sided slope of the active scenario. At kink points, the relevant object is a directional derivative or subgradient, which is handled in the Appendix.

4.2.4 Hard trimming and margin cliffs

To illustrate how hardening a risk input can interact with continuity, consider trimming, which discards sampled returns whose absolute size exceeds a cutoff:

$$\Gamma_t^{\text{trim}} = \sum_{j=0}^{m-1} (P_{t-j} - P_{t-j-1})^2 \cdot \mathbf{1}\{|P_{t-j} - P_{t-j-1}| \leq c\}, \quad c > 0. \quad (22)$$

A return just below c is counted (contributing close to c^2), while a return just above c is dropped (contributing 0). This creates a discrete change in the risk input, and therefore a discrete change in posted margin under any strictly increasing margin mapping.

Proposition 2 (Hard trimming creates discontinuous margin updates). Fix $c > 0$ and define Γ_t^{trim} by (22). Assume g is strictly increasing and continuous. Holding $(P_{t-m}, \dots, P_{t-1})$ fixed, the mapping $P_t \mapsto M_{t+1} = g(\Gamma_t^{\text{trim}})$ is discontinuous at any state with $|R_t| = c$, where $R_t = P_t - P_{t-1}$.

Proof. Holding $(P_{t-m}, \dots, P_{t-1})$ fixed, write $\Gamma_t^{\text{trim}} = A + (P_t - P_{t-1})^2 \mathbf{1}\{|P_t - P_{t-1}| \leq c\}$ for a constant A . At $|P_t - P_{t-1}| = c$, perturbing P_t slightly toward the interior makes the indicator equal to 1 (adding about c^2), while perturbing slightly outward makes it equal to 0 (adding 0). Hence Γ_t^{trim} jumps by c^2 at the cutoff, and by strict monotonicity of g , so does M_{t+1} . \square

Winsorizing (capping the squared return at c^2 rather than dropping it) restores continuity, but creates locally flat regions in which marginal sensitivity is zero once the cap binds. These examples illustrate the design trade-off: efforts to reduce smooth marginal manipulability tend to replace slopes with kinks, caps, or cliffs, and those features interact directly with continuity and risk sensitivity.¹¹

4.3 Strategic constrained traders

The baseline model takes the constrained sector’s target exposure as predetermined at the start of the pricing window. Appendix B relaxes this. Constrained traders are forward-looking: they choose targets anticipating that within-window trading can move the sampled mark, raise the posted requirement $M_{t+1} = g(\Gamma_t)$, and trigger forced deleveraging with price impact. The appendix therefore studies a two-sided subgame-perfect equilibrium in which (i) constrained traders choose targets as a function of the state S_t , and (ii) the strategic trader chooses the trigger size (equivalently a_t or q) in response. The only change to the main mechanism is that equilibrium targeting alters the local forced-flow slope. In the binding states, the relevant object is an effective liquidation elasticity

$$c_X^{\text{eff}}(S_t) = -\frac{\partial X}{\partial M}(M_t, \bar{x}_t^*(S_t)),$$

which is weakly smaller than the non-strategic benchmark because forward-looking traders scale back exposures in states where the update is more susceptible to manipulation. The amplification condition is therefore modified by replacing c_X with $c_X^{\text{eff}}(S_t)$.

Crucially, forward-looking behavior can dampen the loop but cannot eliminate it whenever constraints bind locally. If the cap binds at the equilibrium target, then $c_X^{\text{eff}}(S_t) > 0$, and the same two-period trigger-and-reverse logic applies with c_X replaced by $c_X^{\text{eff}}(S_t)$. Eliminating the channel requires $c_X^{\text{eff}}(S_t) = 0$ in the relevant states, which means that endogenous participation collapses or the constraint is kept slack precisely when risk-sensitive margining is intended to be operative. The equilibrium response therefore shifts the problem into an economically meaningful trade-off: robustness to manipulation is obtained through ex ante contraction of constrained-sector exposure and liquidity.¹² From a design perspective, this strengthens the case for limiting marginal pass-through (slope caps) and for pooling-plus-adjustment schedules. Reducing the local sensitivity of posted requirements lowers both the direct profitability of within-window manipulation and the indirect ex-ante withdrawal of constrained-sector risk-bearing capacity highlighted in Appendix B.

¹¹In the mechanism, profitability is governed by a local amplification: the product of (i) sensitivity of the risk input to sampled prices, (ii) pass-through from the input to posted margin, (iii) the margin elasticity of constrained demand, and (iv) price impact. Slope caps (winsorization) and hard thresholds (trimming, stepwise overlays) limit this amplification by flattening local slopes or concentrating adjustment at kinks. We use these examples only to illustrate how robustification interacts with continuity and sensitivity; the optimal choice of overlays is taken up in the design problem.

¹²Empirically, this outside-option response is not hypothetical: in centrally cleared repo, increases in CCP haircuts push safer borrowers toward OTC trading, with a stronger effect for collateral-constrained borrowers [Chebotarev \(2025\)](#).

5 Multi-asset portfolio margining and cross-contract contagion

It is common to set initial margin at the portfolio level rather than contract by contract. The posted requirement is a scalar function of a portfolio risk statistic computed from a vector of marks. When the requirement binds, meeting a margin call can force liquidation in contracts that are not the ones that most efficiently move the portfolio statistic. This wedge does not arise in the single-contract model: a trader may be able to move the portfolio risk input relatively cheaply in one subset of contracts, while the induced deleveraging is concentrated in a different subset. The main point of this section is that portfolio margining generates a cross-contract manipulation channel and a natural notion of cross-contract contagion. The same feedback loop is at work: a trade at t moves transaction prices P_t , which shifts the portfolio risk input Γ_t and hence the posted requirement M_{t+1} . If the requirement binds, the margin call induces forced rebalancing of constrained positions X_{t+1} , and the resulting liquidation order flow feeds back into prices at $t + 1$. In the multi-asset setting, P_t and X_t are vectors, while the margin call is scalar, so the induced liquidation can be concentrated in specific contracts.

5.1 Setup: vector prices, scalar margin, and cross-margining liquidation

There are $K \geq 1$ cleared contracts. At CCP update times, transaction prices and fundamentals are vectors $P_t, V_t \in \mathbb{R}^K$. Constrained traders hold an aggregate vector position $X_t \in \mathbb{R}^K$, with net order flow $\Delta X_t = X_t - X_{t-1}$. A strategic trader submits a trade vector $a_t \in \mathbb{R}^K$ subject to a capacity constraint $\|a_t\|_1 \leq \bar{a}$.

5.1.1 Price formation and profits

Price formation at update times is linear with cross-impact:

$$P_t = V_t + A(\Delta X_t + a_t), \quad (23)$$

where $A \in \mathbb{R}^{K \times K}$ is an impact matrix. Trading profit over a horizon T is

$$\Pi_T(a_{t:t+T}) = - \sum_{s=0}^T a_{t+s}^\top P_{t+s}. \quad (24)$$

A (vector) round trip over horizon T is a trade sequence (a_t, \dots, a_{t+T}) satisfying $\sum_{s=0}^T a_{t+s} = 0$ in \mathbb{R}^K . Fundamentals satisfy the same martingale restriction as in (1), componentwise:

$$\mathbb{E}[V_{t+1} \mid \mathcal{F}_t] = V_t.$$

As in the single-asset model, this removes predictable drift in fundamentals from the profit calculation; it does not remove predictability generated mechanically by margin-induced liquidation.

5.1.2 Margin rule

A scalar portfolio risk statistic is computed from the sampled vector price path and a scalar margin is posted:

$$\Gamma_t = \Gamma(P_{t-m}, \dots, P_t), \quad M_{t+1} = g(\Gamma_t). \quad (25)$$

Write $P_{t-m:t} = (P_{t-m}, \dots, P_t)$ and $P_{t-m:t-1} = (P_{t-m}, \dots, P_{t-1})$. To state local assumptions, we use the multi-asset analog of the public state in Section 2:

$$S_t = (V_t, P_{t-m}, \dots, P_{t-1}, M_t),$$

where each $P_{t-j} \in \mathbb{R}^K$ is a vector of sampled marks.

5.1.3 Liquidation

The new object relative to the single-asset model is how a scalar margin increase maps into a vector liquidation response. We represent the deviation-induced liquidation response at $t + 1$ by the linear map

$$\Delta X_{t+1} = \Delta X_{t+1}^{(0)} - B_t \Delta M_{t+1}, \quad \Delta M_{t+1} = M_{t+1} - M_{t+1}^{(0)}, \quad (26)$$

where $\Delta X_{t+1}^{(0)}$ and $M_{t+1}^{(0)}$ denote the baseline outcomes absent a time- t deviation, and $B_t \in \mathbb{R}^K$ is an \mathcal{F}_t -measurable liquidation direction capturing cross-margining. Thus a higher scalar requirement forces an additional reduction in the constrained portfolio equal to $B_t \Delta M_{t+1}$, and this reduction need not be aligned with the contract(s) most efficient for moving Γ_t .

5.1.4 Assumptions

Assumption 6 (Impact matrix). The matrix A in (23) has strictly positive diagonal entries and is symmetric positive definite.¹³

Assumption 7 (Directional sensitivity of the portfolio risk statistic). There exist a reachable state S_t , constants $c_\Gamma > 0$ and $\delta_\Gamma > 0$, and a unit direction $d_t \in \mathbb{R}^K$ such that, holding $P_{t-m:t-1}$ fixed, for any perturbation ΔP_t with $d_t^\top \Delta P_t \in [0, \delta_\Gamma]$ and small enough that $P_t + \Delta P_t$ remains in the local neighborhood considered,

$$\Gamma(P_{t-m:t-1}, P_t + \Delta P_t) - \Gamma(P_{t-m:t-1}, P_t) \geq c_\Gamma d_t^\top \Delta P_t. \quad (27)$$

Assumption 8 (Local slope of the margin mapping). There exists an interval $I_\Gamma \subset \mathbb{R}_+$ with $\Gamma_t \in \text{int}(I_\Gamma)$ such that g is differentiable on I_Γ and

$$g'(\gamma) \geq c_g > 0 \quad \text{for all } \gamma \in I_\Gamma.$$

¹³Schneider and Lillo (2019) characterize the no-dynamic-arbitrage restriction for cross-impact. If $I = (A + A^\top)/2$ is not positive semidefinite, pick v with $v^\top I v < 0$; under linear impact $P_t = P_t^{(0)} + A a_t$ and zero expected drift, the round trip $a_t = qv$, $a_{t+1} = -qv$ has expected execution cost $q^2 v^\top A v = q^2 v^\top I v < 0$, hence yields arbitrage. We therefore impose the standard no-manipulation restriction on the impact block, so the round trips we identify arise despite, not because of, an arbitrageable impact model.

Assumption 9 (Cross-margining liquidation is nontrivial). At the reachable state in Assumption 7, the liquidation direction in (26) satisfies $\|B_t\|_1 \geq c_X > 0$.

Assumptions 6–9 are the multi-asset counterparts of the three local links used in the single-asset argument. Assumption 6 pins down how a vector order perturbation maps into transaction prices through the cross-impact matrix A , and imposes the standard no-dynamic-arbitrage restriction on the impact block (so round-trip profits do not arise absent the margin-feedback loop). Assumption 7 is the local one-sided bound on how the scalar portfolio risk statistic reacts to an admissible perturbation of the last sampled mark. Assumption 9 is the local statement that a scalar margin increase induces a nontrivial vector liquidation response, summarized by B_t .

5.2 Local bounds in the multi-asset setting

Fix a reachable state satisfying Assumptions 6–9. We consider local deviations so that (27) applies and so that g is evaluated inside I_Γ . The multi-asset argument uses the same slope chain as in the single-asset model, with vectors and cross-impact.

Lemma 6 (A deviation moves the transaction-price vector). *Holding V_t and ΔX_t fixed, perturbing a_t by Δa_t perturbs the time- t transaction price by*

$$\Delta P_t = A \Delta a_t.$$

Proof. Immediate from (23): $P_t = V_t + A(\Delta X_t + a_t)$. □

Lemma 7 (A price perturbation moves the portfolio risk statistic at first order). *Suppose Assumption 7 holds at S_t . Holding $P_{t-m:t-1}$ fixed, any perturbation ΔP_t with $d_t^\top \Delta P_t \in [0, \delta_\Gamma]$ satisfies*

$$\Delta \Gamma_t = \Gamma(P_{t-m:t-1}, P_t + \Delta P_t) - \Gamma(P_{t-m:t-1}, P_t) \geq c_\Gamma d_t^\top \Delta P_t.$$

Proof. This is exactly (27). □

Lemma 8 (Risk sensitivity of g passes through to margin). *Under Assumption 8, for any $\Delta \Gamma_t \geq 0$ such that $\Gamma_t + \Delta \Gamma_t \in I_\Gamma$,*

$$\Delta M_{t+1} = g(\Gamma_t + \Delta \Gamma_t) - g(\Gamma_t) \geq c_g \Delta \Gamma_t.$$

Proof. By the mean value theorem, $g(\Gamma_t + \Delta \Gamma_t) - g(\Gamma_t) = g'(\xi) \Delta \Gamma_t$ for some $\xi \in (\Gamma_t, \Gamma_t + \Delta \Gamma_t) \subset I_\Gamma$. Assumption 8 gives $g'(\xi) \geq c_g$. □

Lemma 9 (Cross-margining liquidation and cross-impact). *Under (26), the deviation-induced additional constrained order flow at $t + 1$ equals $-B_t \Delta M_{t+1}$. Holding V_{t+1} , $\Delta X_{t+1}^{(0)}$, and a_{t+1} fixed, the liquidation-driven component of the time- $(t + 1)$ transaction price is*

$$\Delta P_{t+1}^{\text{liq}} = -AB_t \Delta M_{t+1}.$$

Proof. The first claim is (26). Substituting $\Delta X_{t+1} = \Delta X_{t+1}^{(0)} - B_t \Delta M_{t+1}$ into (23) at $t + 1$ yields

$$P_{t+1} = V_{t+1} + A(\Delta X_{t+1}^{(0)} - B_t \Delta M_{t+1} + a_{t+1}),$$

so, holding V_{t+1} , $\Delta X_{t+1}^{(0)}$, and a_{t+1} fixed, the deviation contributes $-AB_t\Delta M_{t+1}$. \square

5.3 A two-period cross-asset round trip

We focus on the same object as in the single-contract proof: a two-period round trip that isolates profits generated mechanically by the margin-update channel. Consider a trade vector $a_t \in \mathbb{R}^K$ at t and its reversal $a_{t+1} = -a_t$ at $t + 1$, with $\|a_t\|_1 \leq \bar{a}$. As in the single-asset case, we impose a local condition that removes predictable drift in fundamentals and predictable baseline constrained flow, so that any predictability in prices comes from deviation-induced liquidation.

Assumption 10. At S_t , (i) $\Delta X_t = 0$; (ii) letting $\Delta X_{t+1}^{(0)}$ denote constrained order flow at $t + 1$ absent a deviation at t , $\mathbb{E}[\Delta X_{t+1}^{(0)} \mid S_t] = 0$; (iii) $\mathbb{E}[V_{t+1} \mid S_t] = V_t$.

To keep the algebra readable, we summarize a direction $a \in \mathbb{R}^K$ by two scalars:

$$m_t(a) = d_t^\top Aa, \quad l_t(a) = a^\top AB_t.$$

The first quantity, $m_t(a)$, measures how strongly the time- t price move loads on the risk-increasing direction d_t because $\Delta P_t = Aa$ and $d_t^\top \Delta P_t = m_t(a)$. The second quantity, $l_t(a)$, measures exposure to the liquidation-driven price component at $t + 1$, since liquidation moves prices in the direction AB_t .

Lemma 10 (A profitable two-period round trip). *Suppose Assumptions 6–9 and 10 hold, and consider the two-period round trip $a_{t+1} = -a_t$ with $\|a_t\|_1 \leq \bar{a}$. Assume the deviation is local so that $m_t(a_t) \in [0, \delta_\Gamma]$ and $\Gamma_t + \Delta\Gamma_t \in I_\Gamma$. If $l_t(a_t) \leq 0$, then*

$$\mathbb{E}[\Pi_1 \mid S_t] \geq -(c_g c_\Gamma) l_t(a_t) m_t(a_t) - 2a_t^\top Aa_t. \quad (28)$$

Proof. Under $a_{t+1} = -a_t$, two-period trading profit is $\Pi_1 = a_t^\top (P_{t+1} - P_t)$. Using (23) at t and $t + 1$, together with $\Delta X_t = 0$ and $\Delta X_{t+1} = \Delta X_{t+1}^{(0)} - B_t \Delta M_{t+1}$ from (26), we obtain

$$P_t = V_t + Aa_t, \quad P_{t+1} = V_{t+1} + A(\Delta X_{t+1}^{(0)} - B_t \Delta M_{t+1} - a_t),$$

hence

$$P_{t+1} - P_t = (V_{t+1} - V_t) + A\Delta X_{t+1}^{(0)} - AB_t \Delta M_{t+1} - 2Aa_t.$$

Taking $\mathbb{E}[\cdot \mid S_t]$ and applying Assumption 10 yields

$$\mathbb{E}[P_{t+1} - P_t \mid S_t] = -AB_t \Delta M_{t+1} - 2Aa_t,$$

so

$$\mathbb{E}[\Pi_1 \mid S_t] = -(a_t^\top AB_t) \Delta M_{t+1} - 2a_t^\top Aa_t = -l_t(a_t) \Delta M_{t+1} - 2a_t^\top Aa_t.$$

It remains to lower bound ΔM_{t+1} . By Lemma 6, $\Delta P_t = Aa_t$, so $d_t^\top \Delta P_t = m_t(a_t)$. If $m_t(a_t) \in [0, \delta_\Gamma]$, Assumption 7 gives $\Delta\Gamma_t \geq c_\Gamma m_t(a_t)$, and Lemma 8 gives $\Delta M_{t+1} \geq c_g \Delta\Gamma_t \geq c_g c_\Gamma m_t(a_t)$. Substituting into the expression for $\mathbb{E}[\Pi_1 \mid S_t]$ yields (28). \square

Lemma 10 isolates what is genuinely new under portfolio margining. The time- t leg must raise the scalar portfolio risk input, captured by $m_t(a_t) = d_t^\top Aa_t > 0$. But the predictable price component that generates profit at $t + 1$ is liquidation-driven and points in the cross-impact direction AB_t , so the relevant exposure of the round trip is $l_t(a_t) = a_t^\top AB_t$. Under contract-by-contract margining these objects collapse to the same sign choice; under portfolio margining they generally decouple because the risk-increasing direction d_t and the forced-deleveraging direction B_t need not align. This creates a trigger-versus-harvest wedge: a trader can move the scalar charge efficiently using one set of contracts while taking offsetting exposure in the contracts that are unwound when the call binds, with the strength governed by the alignment term $-d_t^\top AB_t$. This is distinct from standard predatory-trading/fire-sale mechanisms.

5.4 A sufficient amplification condition

Lemma 10 gives a direction-dependent lower bound. For interpretation, it is useful to have a simple sufficient condition that guarantees existence of at least one admissible two-period round trip with strictly positive conditional expected profit. The next result does this by selecting a concrete direction, namely the cross-margining liquidation direction B_t , and checking when the margin channel dominates the trader's own round-trip impact loss.

Theorem 3 (A sufficient condition for a profitable two-period round trip). *Suppose Assumptions 6–10 hold and that there exists a liquidation direction $B_t \in \mathbb{R}^K$ with $\|B_t\|_1 = 1$ and $d_t^\top AB_t < 0$. Define*

$$H_t = (B_t^\top AB_t) \left(-(c_g c_\Gamma) d_t^\top AB_t - 2 \right). \quad (29)$$

If $H_t > 0$ and $q > 0$ is admissible (so that the locality bounds apply), then the two-period round trip $a_t = -qB_t$, $a_{t+1} = qB_t$ satisfies

$$\mathbb{E}[\Pi_1 \mid S_t] \geq q^2 H_t > 0.$$

In particular, whenever locality permits arbitrarily small $q > 0$, profitable round trips exist for arbitrarily small admissible trigger sizes.

Proof. Let $a_t = -qB_t$ and $a_{t+1} = qB_t$. Then $l_t(a_t) = a_t^\top AB_t = -qB_t^\top AB_t \leq 0$ and $m_t(a_t) = d_t^\top Aa_t = -q d_t^\top AB_t > 0$. For admissible q , we have $m_t(a_t) \in [0, \delta_\Gamma]$ and $\Gamma_t + \Delta\Gamma_t \in I_\Gamma$, so Lemma 10 applies:

$$\mathbb{E}[\Pi_1 \mid S_t] \geq -(c_g c_\Gamma) l_t(a_t) m_t(a_t) - 2a_t^\top Aa_t = q^2 H_t.$$

If $H_t > 0$, this bound is strictly positive for every admissible $q > 0$. When $K = 1$ and $A = \alpha > 0$, taking $d_t = -1$ and $B_t = c_X$ reduces $H_t > 0$ to the scalar amplification condition. \square

Theorem 3 is only a sufficient condition. We choose B_t because it has a clear meaning: it is the direction in which the constrained sector is forced to sell when a portfolio margin call arrives. The economic message is that portfolio margining separates two things that coincide in the single-contract model. A trader can trigger the margin call using the contracts that move the portfolio risk number most cheaply, but the forced selling that follows can be concentrated in different contracts. By loading

on those contracts, the trader can earn predictable profits from the margin-induced flow at $t + 1$, even though fundamentals themselves have no predictable drift.

5.5 Multi-asset impossibility

Theorem 4 (Impossibility under portfolio margining and cross-contract contagion). *Fix a reachable state S_t satisfying the local assumptions in Section 5. Let $\tilde{q}_{\max} > 0$ denote the maximal trigger size allowed by locality and capacity in the multi-asset setting. If $H_t > 0$ as defined in (29), then for every admissible $q \in (0, \tilde{q}_{\max}]$ the two-period round trip $a_t = -qB_t$, $a_{t+1} = qB_t$ has strictly positive conditional expected profit:*

$$\mathbb{E}[\Pi_1 \mid S_t] \geq q^2 H_t > 0.$$

In particular, portfolio margining can generate profitable round trips for arbitrarily small admissible trigger sizes. Consequently, risk sensitivity (Definition 2), round-trip manipulation-proofness (Definition 3), and liquidity continuity (Definition 4) cannot hold simultaneously at S_t .

Proof. Fix any admissible $q \in (0, \tilde{q}_{\max}]$ and take $a_t = -qB_t$, $a_{t+1} = qB_t$. Under the local conditions defining H_t , the bound in Theorem 3 applies and gives $\mathbb{E}[\Pi_1 \mid S_t] \geq q^2 H_t$. If $H_t > 0$, expected profit is strictly positive for every admissible $q > 0$, which violates round-trip manipulation-proofness (Definition 3) at S_t . The incompatibility with risk sensitivity and liquidity continuity follows by the same contrapositive logic as in Theorem 1. \square

Portfolio margining turns the mechanism into a genuinely cross-contract channel. A vector of marks is mapped into a single portfolio charge, but a binding charge forces liquidation in a particular direction B_t , which need not coincide with the contracts that most efficiently move the portfolio statistic. A trader can therefore trigger a higher portfolio margin using one set of contracts and profit from the predictable forced flow in another set of contracts, with cross-impact transmitting that flow into prices across the cleared complex even when fundamentals are martingales.¹⁴

5.6 What is new relative to the single-contract case

Portfolio margining creates a wedge that is absent in the single-contract model. In the scalar case, the same contract both determines the risk input that sets next-period margin and absorbs the forced liquidation when margin tightens. Under portfolio margining, the CCP maps a vector of marks into a single scalar charge, but a binding charge can force liquidation in a different set of contracts. This separation is the source of a genuinely new cross-contract channel.

The trade at t moves marks by $\Delta P_t = Aa_t$, so the first-order effect on the portfolio risk input is governed by

$$d_t^\top \Delta P_t = d_t^\top Aa_t,$$

the direction that increases the portfolio statistic (the trigger). The resulting margin increase forces additional constrained selling in direction B_t , which moves next-period prices through cross-impact

¹⁴As an example, this coupling is operational in SPAN-like systems, where combined-commodity offsets and portfolio netting link contracts through a common scalar charge (CME Group, 2019b,c).

in direction AB_t . The trader's exposure to this liquidation-driven price component is governed by

$$a_t^\top AB_t,$$

the harvest. When $K = 1$ these objects collapse to the same sign choice. When $K > 1$ they decouple: it can be feasible to choose a_t so that the trade at t raises the scalar portfolio charge ($d_t^\top Aa_t > 0$) while the position is short the contracts that are sold to meet the margin call ($a_t^\top AB_t < 0$). Economically, the trader triggers the margin call in the cheapest contracts and harvests the forced-flow price pressure in different contracts.

This trigger-versus-harvest separation is not an artifact of a two-period construction. In a dynamic treatment of price-based constraints, [Nicolai \(2026\)](#) shows that optimal finite-horizon round trips can be confined to the span of two portfolios: a trigger portfolio that most efficiently moves the marked statistic and a liquidation portfolio that captures the forced-flow direction. In our notation these are the directions governing $d_t^\top Aa_t$ and AB_t . The portfolio case can make manipulation easier. Manipulation-proofness must rule out all admissible vector round trips. As K grows, the space of directions that can simultaneously raise the scalar risk input and load negatively on liquidation expands, so it is enough that a single direction is profitable at a reachable state for manipulation-proofness to fail. At the same time, the strength of the effect is governed by an economically interpretable alignment term: how strongly the risk-increasing direction and the liquidation-driven price direction line up. In the sufficient condition, the key scalar is $-d_t^\top AB_t$.

Finally, portfolio margining implies an empirical signature that does not arise naturally under contract-by-contract margining. The contracts where trading pressure appears at t (those that most efficiently move the portfolio risk input) need not be the same contracts where liquidity deteriorates at $t + 1$ (those along the forced-liquidation direction). Churn and subsequent stress can show up in different legs of the cleared complex. This is a direct implication of mapping a scalar portfolio charge into a vector liquidation response, and it is a sharp diagnostic of cross-contract contagion under portfolio margining. Empirically, there are episodes in which collateral calls on derivatives portfolios have translated into concentrated sales and liquidity dislocations in related cash markets, including the UK gilt episode (Sep–Oct 2022) ([Pinter, 2023](#)) and the European energy derivatives episode in 2022 ([Avalos et al., 2023](#)).

6 Optimal Margin Design

This section studies how a designer should map a price-based risk input into posted initial margin:

$$M_{t+1} = g(\Gamma_t),$$

where Γ_t is computed from sampled transaction prices (or marks) as specified by the rulebook. The earlier results imply that, in binding states with price impact, the schedule is a design object: its local pass-through affects both margin-induced flows and the incentives created by using tradable marks as inputs. Accordingly, we choose g to balance loss coverage against collateral costs, subject to an implementability restriction that limits short-horizon pass-through. We impose this restriction

directly as a slope cap on g (equation (34) below), which is a sufficient and tractable constraint in the baseline. Appendix C derives all the results.

6.1 Loss coverage

Let $D_{t+1} \geq 0$ denote the next-period loss that initial margin at $t + 1$ is intended to cover (default loss, close-out loss, liquidation loss, or settlement exposure). To connect losses to the observable risk input, we assume a representation:

$$D_{t+1} = \sqrt{\Gamma_t} Y_{t+1}, \quad (30)$$

where $Y_{t+1} \geq 0$ is a standardized shock whose conditional distribution does not depend on the level of Γ_t (and, on the benchmark path, does not depend on the choice of g). This captures the core operational interpretation of many CCP risk numbers: Γ_t is a variance-type state, so losses scale with $\sqrt{\Gamma_t}$ in the same way that VaR/ES scale with volatility in common parametric implementations. A direct consequence is square-root scaling of conditional quantiles:

$$Q_p(D_{t+1} \mid \Gamma_t = \gamma) = \sqrt{\gamma} Q_p(Y_{t+1}), \quad (31)$$

so any quantile-based target schedule is automatically increasing in γ .

6.2 The designer's objective

We model the margin designer as choosing g to minimize a stationary one-period criterion that trades off three components:

$$J(g) = \mathbb{E}[\ell(D_{t+1} - g(\Gamma_t)) + \lambda \mathcal{C}(X_{t+1} - X_t) + \kappa_M g(\Gamma_t)]. \quad (32)$$

The first term is a coverage loss: ℓ penalizes shortfall of posted margin relative to realized losses. The second term is a market-functioning/procyclicality cost: when constraints bind, a margin increase forces deleveraging, and we penalize the induced forced flow with a convex function \mathcal{C} . The third term is an opportunity cost of posted margin, $\kappa_M > 0$, capturing collateral, funding, and participation costs. The criterion in (32) is intended as a tractable implementation of PFMI-style coverage and stability goals, not as a full general-equilibrium model of CCP governance.¹⁵ The coverage-loss term maps to risk-based margin intended to cover potential future exposure; the market-functioning term maps to mitigating procyclicality and avoiding liquidity cliffs from large, mechanically induced calls; and the collateral-cost term acknowledges collateral liquidity and opportunity costs for clearing members.¹⁶ With the impossibility-implied slope cap in place, the optimal degree of pooling is pinned down by these PFMI-aligned trade-offs together with implementability. A key point for de-

¹⁵For a carefully developed contracting analysis that endogenizes the default waterfall, CCP compensation, and equilibrium capital contributions from monitoring incentives and collateral costs, see Kuong and Maurin (2024). For a complementary framework emphasizing the incentive role of collateral requirements and the IM versus default-fund composition trade-off, see Wang et al. (2022).

¹⁶See, for instance, Committee on Payments and Market Infrastructures and International Organization of Securities Commissions (2012); Basel Committee on Banking Supervision and Committee on Payments and Market Infrastructures and Board of the International Organization of Securities Commissions (2022); European Systemic Risk Board (2020); Bank of England (2021)

sign is that the procyclicality channel is inherently tied to margin updates. Locally, in binding states, constrained exposure is margin-sensitive:

$$X_{t+1} - X_t \approx -c_X (M_{t+1} - M_t), \quad c_X > 0.$$

Thus, for given Γ -dynamics, a steeper mapping g mechanically produces larger margin updates and larger forced-flow. To make this trade-off transparent, Appendix C.2 adopts a smooth approximation that converts the update-cost term into a linear functional of the local slope $g'(\cdot)$. The steps are the following: take $\mathcal{C}(z) = |z|$ (an absolute value penalty in the magnitude of forced flow), linearize $g(\Gamma_{t+1}) - g(\Gamma_t) \approx g'(\Gamma_t) (\Gamma_{t+1} - \Gamma_t)$ for small one-step input changes, and replace $|\Gamma_{t+1} - \Gamma_t|$ by its conditional mean

$$\delta(\gamma) = \mathbb{E}[|\Gamma_{t+1} - \Gamma_t| \mid \Gamma_t = \gamma].$$

Under stationarity, the resulting design problem can be written directly as

$$\min_{g \in \tilde{\mathcal{G}}} \int_0^\infty f(\gamma) \left(\phi_\gamma(g(\gamma)) + \kappa_M g(\gamma) \right) d\gamma + \lambda c_X \int_0^\infty f(\gamma) \delta(\gamma) g'(\gamma) d\gamma, \quad (33)$$

where $\phi_\gamma(m) = \mathbb{E}[\ell(D_{t+1} - m) \mid \Gamma_t = \gamma]$ is the conditional expected coverage loss. The first integral is the state-by-state coverage and level-cost objective. The second integral prices short-horizon pass-through: in states where the risk input is more volatile (high $\delta(\gamma)$), a higher local slope $g'(\gamma)$ produces larger expected margin updates and hence larger forced-flow impulses. Appendix C.2 derives (33) and solves it.

When $\lambda = 0$ and $\ell(x) = (x)_+$, the problem separates state by state, and the first-best margin is a conditional quantile. In particular, for each γ , the statewise minimizer satisfies

$$\Pr(D_{t+1} > g^*(\gamma) \mid \Gamma_t = \gamma) = \kappa_M,$$

so $g^*(\gamma) = Q_{1-\kappa_M}(D_{t+1} \mid \Gamma_t = \gamma)$, which inherits square-root scaling under (30). This benchmark and its scaling implication are in Appendix C.

6.3 Implementability as a slope cap on $g'(\cdot)$ and the shape of the solution

The impossibility mechanism implies an implementability restriction: in binding states, local pass-through from the price-based input into posted margin cannot be arbitrarily steep without creating profitable trigger-and-reverse round trips. In the realized-variance and linear-impact benchmark, Corollary 1 delivers a pointwise envelope of the form $g'(r) \leq C/\sqrt{r}$. A convenient global sufficient restriction is the cap

$$0 \leq g(\gamma_2) - g(\gamma_1) \leq \bar{s} (\sqrt{\gamma_2} - \sqrt{\gamma_1}) \quad \text{for all } \gamma_2 \geq \gamma_1, \quad (34)$$

which is equivalent (where differentiable) to the slope constraint

$$0 \leq g'(\gamma) \leq \frac{\bar{s}}{2\sqrt{\gamma}} \quad \text{for } \gamma > 0.$$

In that same benchmark, a sufficient implementable choice consistent with Corollary 1 is

$$\bar{s} = \frac{2\alpha + \kappa}{\alpha^2 c_X}, \quad (35)$$

with $\kappa = 0$ in the baseline linear-impact model. Two points are worth emphasizing.

First, slope caps are not merely anti-procyclicality overlays. They are also an incentive-compatibility device: they directly limit the local loop gain that a trigger trade can create through margin updates and forced liquidation. Moreover, when constrained traders are forward-looking, steep pass-through creates an additional ex ante cost: it raises perceived manipulation-driven margin-jump risk, which reduces equilibrium constrained participation and lowers the effective liquidation elasticity (c_X^{eff}). This dampening is stabilizing but comes from reduced leveraged participation and reduced liquidity provision. Appendix B.5 formalizes this participation trade-off and shows why limiting $g'(\cdot)$ is complementary to maintaining robust participation in stress.

Second, with the smooth objective, the slope cap also disciplines the shape of the optimal schedule. Because the update-cost term in (33) is linear in $g'(\cdot)$, the optimizer generically selects corner solutions for the local slope. Appendix C.2 shows that g consists of three interpretable regimes. In pooling regions the schedule is locally flat, $g'(\gamma) = 0$, so small movements in the risk input do not change posted requirements; economically, this implements a buffer or floor that mutes marginal pass-through in calm or moderately stressed states. In an intermediate tracking region the slope is interior, $0 < g'(\gamma) < \bar{s}/(2\sqrt{\gamma})$, and the schedule moves with the statewise coverage target implied by the effective marginal cost of margin (coverage versus collateral costs plus the local shadow price of procyclicality). In stressed states the implementability constraint binds and the optimizer exhibits slope saturation, $g'(\gamma) = \bar{s}/(2\sqrt{\gamma})$, so margin increases as fast as the cap permits; under the volatility-scale cap, this implies $g(\gamma)$ is affine in $\sqrt{\gamma}$ on saturated segments. This pooling/tracking/saturation structure is the continuous-state analogue of standard CCP rulebook tools: buffers and floors correspond to pooling, intermediate regimes correspond to target-like behavior, and limited pass-through in stress corresponds to saturation. Figure 1 schematically illustrates the resulting optimal shape.

6.4 Portfolio margining: netting versus cross-impact and the trigger-liquidation wedge

When margins are set at the portfolio level, a vector of marks is mapped into a scalar risk statistic; a scalar requirement $M_{t+1} = g(\Gamma_t)$ is posted, and the market meets a binding call through a portfolio deleveraging response. This changes both the market-functioning costs of a margin update and the relevant implementability restriction. With $K \geq 2$ contracts and symmetric positive definite cross-impact matrix A , transaction prices satisfy $P_t = V_t + A Q_t$. A binding scalar margin increase induces forced liquidation in a direction $B_t \in \mathbb{R}^K$, so the update-driven forced flow is proportional to the scalar margin change:

$$v_{t+1} \approx -B_t(g(\Gamma_t) - g(\Gamma_{t-1})), \quad \Delta P_{t+1}^{\text{liq}} \approx -AB_t(g(\Gamma_t) - g(\Gamma_{t-1})).$$

Thus, portfolio margining makes the procyclicality channel explicitly vector-valued: a scalar call generates a portfolio liquidation, and cross-impact transmits that liquidation into several prices. Ap-

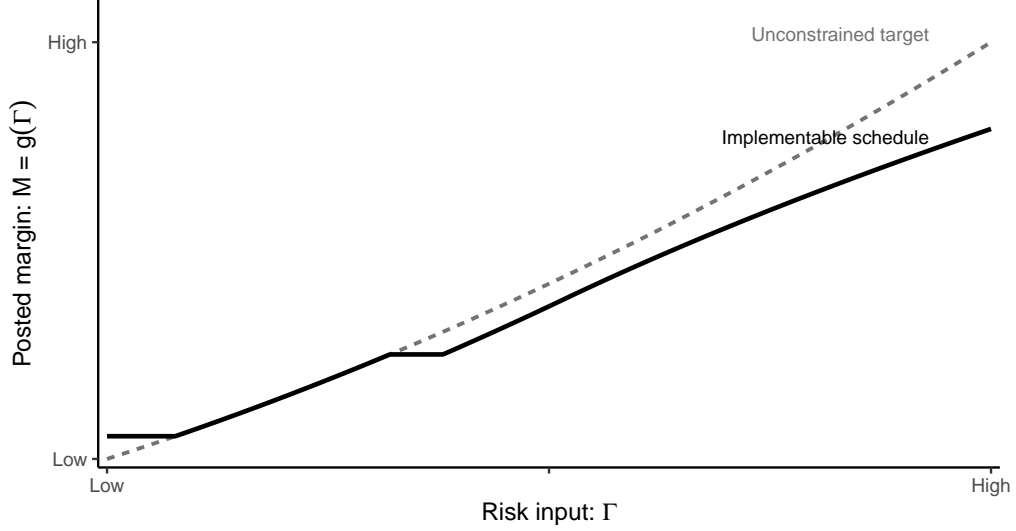


Figure 1 : Optimal Margin Rule

The figure displays the optimal margin schedule $M_{t+1} = g(\Gamma_t)$. The dotted grey line shows the unconstrained schedule, while the solid black line shows the implementable schedule obtained after imposing the incentive-compatibility constraint as a pointwise cap on $g'(\cdot)$.

pendix C.3 applies the same smooth approximation used in the single-asset case to obtain a continuous objective of the same form as (33), but with the update-cost weight now reflecting portfolio liquidation intensity and cross-impact. Economically, the single-asset coefficient $\delta(\gamma)$ is replaced by an effective state-dependent expected magnitude of a margin update times the marginal cost of the induced liquidation vector (which depends on B_t and on AB_t). The solution structure remains pooling/tracking/slope saturation, but the relevant slopes and shadow prices are portfolio objects.

Portfolio margining creates a wedge absent in the single-contract case: the contracts that move the scalar risk input most efficiently (the trigger direction) need not coincide with the contracts liquidated when a call binds. The impossibility result in Section 5 shows that this decoupling generates a genuinely cross-contract manipulation channel. For design, the key implication is that the manipulation-proof slope bound depends on an alignment coefficient that measures how strongly the liquidation-driven price-pressure direction loads on the risk-increasing direction in mark space. In the sufficient restriction derived in Appendix C.3, the relevant scalar is

$$\psi_t = \max\{0, -d_t^\top AB_t\},$$

where d_t is the local risk-increasing direction for the portfolio statistic. Larger ψ_t tightens the admissible bound on $g'(\cdot)$, which mechanically expands pooling regions and flattens adjustment regions under the design objective.

The operational takeaway is that portfolio margining creates a tension between levels and pass-through. On the one hand, aggregation and netting typically reduce the level of required margin for a given set of exposures by recognizing offsets and diversification. On the other hand, when

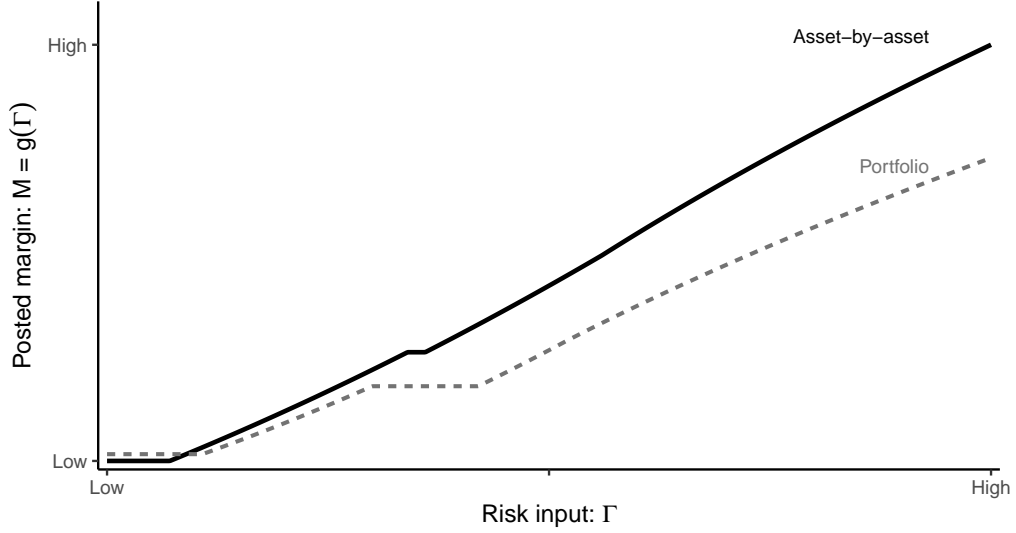


Figure 2 : Asset-by-Asset vs Portfolio Margining

The figure displays margin schedules under asset-level margining (solid black line) and portfolio-level margining (dotted grey line).

liquidation is concentrated and cross-impact is material, the marginal market-functioning cost of short-horizon updates rises and the implementable pass-through constraint can tighten through the portfolio multiplier ψ_t . In the design problem this pushes the optimal schedule toward more pooling and flatter adjustment, unless the CCP compensates by raising a floor or buffer to preserve coverage.

Appendix C.3 provides a two-asset illustration in which cross-impact β enters the portfolio slope cap through $\psi_t = b\beta$, while the asset-by-asset cap depends primarily on diagonal impact. Figure 2 illustrates the slope effect: for the same scalar risk input, the portfolio rule is less locally responsive because the implementability cap is tighter.

6.5 Why jump discontinuities do not help

A natural instinct is to replace steep slopes by discrete steps (thresholds, rounding, stepwise additions). In this environment, upward jumps do not solve the incentive problem; they relocate it to threshold-crossing deviations. If the intended input is arbitrarily close below a jump threshold, an arbitrarily small trade can cross the threshold, trigger a discrete increase in the next call, and thereby induce a discrete liquidation event. When constraints bind, this creates a predictable price component that a short-horizon round trip can monetize.

Appendix C.5 provides the formal statements. In the single-asset case, an upward jump of size ΔM creates a profitable threshold-crossing deviation whenever the induced liquidation is large enough relative to execution costs. In the portfolio case, the same logic applies with cross-impact: a jump implies a vector forced liquidation $v = -B_t \Delta M$ and liquidation-driven price pressure $-AB_t \Delta M$, so a trader can choose a round-trip direction that loads on that predictable component. The broader design implication is that bounded-slope adjustment regions dominate jumps as a robustness device:

they remove vanishing-cost threshold manipulation by spreading adjustment over a finite width of γ .

6.6 Interpretation and implementation

The pooling-plus-adjustment structure delivers a structural interpretation of common CCP anti-procyclicality and robustness tools. Pooling regions correspond to states in which an overlay binds (floors, buffers, minimum charges), so small perturbations of sampled marks do not change posted margin at first order. Tracking regions correspond to intermediate regimes where the CCP can follow the coverage target implied by its objective. Slope-saturated regions correspond to stress regimes where limited pass-through is binding: posted margin responds to risk numbers, but only up to the implementable cap implied by the incentive restriction. Operational discretization (rounding conventions, bucketed add-ons, threshold triggers for intraday calls) can be read as a coarse implementation of these regions. In the model, these features arise endogenously from the fact that the CCP is using tradable transaction-based inputs to set a binding requirement that feeds back into prices through forced liquidation. The impossibility result therefore provides a structural rationale for why CCP rulebooks place emphasis not only on coverage levels but also on stability and bounded pass-through of margin updates.

7 Conclusions

This paper studies price-based risk constraints: disclosed rulebooks that map sampled transaction prices into a risk input and then into a binding requirement, $M_{t+1} = g(\Gamma_t)$. In binding regimes with price impact, the update rule is part of the trading environment, because within-window trades can move the sampled marks that determine next-period feasibility and the associated forced flow.

The core result is a local, constructive impossibility. In neighborhoods of reachable states in which a nontrivial constrained sector is already locally binding, local risk sensitivity, liquidity continuity, and round-trip manipulation-proofness cannot generally be achieved simultaneously (Theorem 1). If, under a fixed rule, the system can reach with positive probability a binding high-amplification state (and empirically such states do occur), then no designer can choose the rule to be locally risk-sensitive, liquidity-continuous, and round-trip manipulation-proof. The proof constructs a trigger-and-reverse deviation: an arbitrarily small trade tilts a sampled mark, shifts M_{t+1} at first order through $g(\Gamma_t)$, induces predictable forced liquidation when the requirement binds, and is reversed into that liquidation-driven price pressure. Profitability is governed by the same local amplification chain that links marks to the statistic, the statistic to the requirement, the requirement to forced sales, and forced sales to prices; equations (14) and (15) provide a quantitative lower bound on the gain.

Three scope clarifications preempt common misreadings. First, the result is local: under the gain condition, expected inventory-adjusted profits are strictly positive for every admissible trigger size in a neighborhood, so the mechanism does not rely on large trades or large price moves. Second, the result is about binding regimes: it does not rely on making a slack constraint bind, but on marginal tightening in states where the requirement already binds. Third, the predictable component is rule-induced rather than fundamental: the argument does not require predictable drift or persistent mis-

pricing, and it applies even when the impact block is manipulation-free absent feedback.

Read as a design restriction, the impossibility delivers sharp empirical content. Corollary 1 converts the incentive constraint into an implied bound on short-horizon pass-through from transaction-based inputs into posted requirements. Section 3.6 maps the same amplification objects into observable magnitudes using a conservative link from participation to within-window mark distortion together with empirically grounded liquidation elasticities and realistic one-update requirement changes, yielding economically meaningful return-on-margin lower bounds precisely in binding states. A distinct implication concerns state dependence: for common implementations in which VaR/ES is pinned down by an estimated volatility, the marginal sensitivity of Γ_t to a single sampled mark can be highest after quiet spells, so conditional on binding, tranquil states can be locally most susceptible even when standard diagnostics report low risk (Section 4.2).

The tension survives realistic implementation details. Continuous trading between discrete measurement and reset times does not remove the channel because it is pinned to discrete sampling and enforcement. Allowing constrained traders to be forward-looking changes only the effective forced-flow elasticity; it cannot restore joint feasibility in any state in which constraints remain locally binding. Non-smooth overlays can reduce smooth marginal pass-through away from thresholds, but they also concentrate adjustment at kinks and can shift incentives toward threshold-crossing deviations.

Portfolio margining strengthens the mechanism and makes it cross-contract. When a scalar requirement is computed from a vector of marks, the contracts that most efficiently move the portfolio statistic need not coincide with the contracts liquidated to meet a binding call. Section 5 formalizes this trigger-versus-harvest wedge and shows that it generates cross-contract manipulation and spillovers under cross-impact (Theorem 4), with a sharp signature: within-window trading pressure can appear in different legs than the post-update liquidity stress associated with forced liquidation.

Finally, the paper uses the implied implementability restriction to discipline design. Section 6 treats g as a choice variable and imposes a slope cap as a tractable sufficient restriction limiting short-horizon pass-through. The resulting optimal schedules exhibit pooling and deliberately muted adjustment regions, with slope saturation in stress states, providing a structural rationale for bounded pass-through, smoothing, buffers, and related anti-procyclicality tools as incentive-compatibility devices rather than purely statistical overlays.

Several directions follow directly. Empirical tests need not start with CCPs, since CCP rulebooks already deploy anti-procyclicality tools that mute short-horizon pass-through and can compress variation in local slopes; this can shift incentives toward kinks and thresholds and complicate detection even when the feedback is present. A cleaner laboratory is mandated investment rules with transparent mechanical rebalancing; Nicolai (2026) applies the same logic to volatility-managed portfolios and converts the amplification diagnostic into a liquidity-scaled capacity bound. The multi-asset results motivate event studies that separate trigger directions from liquidation directions under portfolio systems. On the design side, we take the statistic as given; a natural complement is statistic design, namely constructions of Γ_t that minimize directional sensitivity to tradable marks subject to coverage and robustness. More broadly, embedding endogenous participation, clearing choices, and equilibrium liquidity provision into the design problem would connect the implementability restriction to welfare and to the composition effects highlighted by strategic responses. Because the

mechanism is about price-based constraints rather than any single institution, the same diagnostic and design logic applies wherever disclosed rulebooks map recent prices into binding feasibility.

References

- Adrian, Tobias and Hyun Song Shin, “Liquidity and leverage,” *Journal of Financial Intermediation*, 2010, 19 (3), 418–437.
- and —, “Procyclical Leverage and Value-at-Risk,” *The Review of Financial Studies*, 2014, 27 (2), 373–403.
- Aldasoro, Inaki, Torsten Ehlers, and Egemen Eren, “Liquid assets at CCPs and systemic liquidity risks,” *BIS Quarterly Review*, December 2023.
- Allen, Franklin and Douglas Gale, “Stock-price manipulation,” *Review of Financial Studies*, 1992, 5 (3), 503–529.
- Aramonte, Sirio, Andreas Schrimpf, and Hyun Song Shin, “Margins, Debt Capacity, and Systemic Risk,” BIS Working Papers 1121, Bank for International Settlements September 2023.
- Avalos, Fernando and Vladyslav Sushko, “Margin Leverage and Vulnerabilities in US Treasury Futures,” *BIS Quarterly Review*, September 2023, Box A September 2023.
- , Wenqian Huang, and Kevin Tracol, “Margins and Liquidity in European Energy Markets in 2022,” BIS Bulletin 77, Bank for International Settlements September 2023.
- Bank of England, “A CBA of APC: analysing approaches to procyclicality reduction in CCP initial margin models,” Staff Working Paper 950, Bank of England 2021.
- Barth, Daniel and Jay Kahn, “Basis Trades and Treasury Market Illiquidity,” Brief Series 20-01, Office of Financial Research, U.S. Department of the Treasury July 2020.
- Basak, Suleyman and Alexander Shapiro, “Value-at-Risk-Based Risk Management: Optimal Policies and Asset Prices,” *The Review of Financial Studies*, 2001, 14 (2), 371–405.
- Basel Committee on Banking Supervision, “Transparency and responsiveness of initial margin in centrally cleared markets: review and policy proposals,” Technical Report d568, Bank for International Settlements 2022.
- Basel Committee on Banking Supervision and Committee on Payments and Market Infrastructures and Board of the International Organization of Securities Commissions, “Review of Margining Practices,” Technical Report, Bank for International Settlements and IOSCO September 2022. Available at: <https://www.iosco.org/library/pubdocs/pdf/IOSCOPD714.pdf>.
- Biais, Bruno, Florian Heider, and Marie Hoerova, “Risk-sharing or risk-taking? Counterparty risk, incentives, and margins,” *Journal of Finance*, 2016, 71 (4), 1669–1698.
- Bignon, Vincent and Guillaume Vuillemeys, “The Failure of a Clearinghouse: Empirical Evidence,” *Review of Finance*, 2020, 24 (1), 99–128.

- Borio, Claudio and Mathias Drehmann, “Towards an operational framework for financial stability:” fuzzy” measurement and its consequences,” *Documentos de Trabajo (Banco Central de Chile)*, 2009, (544), 1.
- Brunnermeier, Markus K. and Lasse Heje Pedersen, “Market Liquidity and Funding Liquidity,” *Review of Financial Studies*, 2009, 22 (6), 2201–2238.
- and Yuliy Sannikov, “A Macroeconomic Model with a Financial Sector,” *American Economic Review*, 2014, 104 (2), 379–421.
- Chebotarev, Dmitry, “Adverse Selection Effect of CCP Haircuts,” March 2025. Working paper, Indiana University Kelley School of Business.
- CME Group, “CME Group to Launch Next Generation of CME SPAN Margin Methodology,” May 2019.
- , “CME SPAN Methodology Overview,” 2019.
- , “SPAN Methodology,” Technical Report, CME Group March 2019.
- , “SPAN Reference Documents,” 2026.
- Committee on Payments and Market Infrastructures and International Organization of Securities Commissions, “Principles for Financial Market Infrastructures,” Technical Report, Bank for International Settlements and IOSCO April 2012.
- Commodity Futures Trading Commission, “17 CFR 38.200 (Core Principle 3): Contracts Not Readily Susceptible to Manipulation,” <https://www.law.cornell.edu/cfr/text/17/38.200> 2026. Electronic Code of Federal Regulations (e-CFR).
- , “Futures Glossary: Settlement Price,” 2026.
- Cont, Rama and Seyyedmostafa Ghamami, “Skin in the game: risk analysis of central counterparties,” *Journal of Financial Market Infrastructures*, 2025.
- and Thomas Kokholm, “Central clearing of OTC derivatives: Bilateral vs multilateral netting,” *Statistics and Risk Modeling*, 2014, 31 (1), 3–22.
- , Romain Deguest, and Giacomo Scandolo, “Robustness and sensitivity of risk measures,” *Quantitative Finance*, 2010, 10 (6), 593–606.
- Coval, Joshua and Erik Stafford, “Asset fire sales (and purchases) in equity markets,” *Journal of Financial Economics*, 2007, 86 (2), 479–512.
- Danielsson, Jón, Hyun Song Shin, and Jean-Pierre Zigrand, “The impact of risk regulation on price dynamics,” *Journal of Banking & Finance*, 2004, 28 (5), 1069–1087.
- Duarte, Fernando and Thomas M. Eisenbach, “Fire-Sale Spillovers and Systemic Risk,” Staff Reports 645, Federal Reserve Bank of New York October 2013.

- Duffie, Darrell, “Still the World’s Safe Haven? Redesigning the U.S. Treasury Market After the COVID-19 Crisis,” Hutchins Center Working Paper 62, Brookings Institution June 2020.
- and Haoxiang Zhu, “Does a central clearing counterparty reduce counterparty risk?,” *Review of Asset Pricing Studies*, 2011, 1 (1), 74–95.
- , Martin Scheicher, and Guillaume Vuillemeys, “Central Clearing and Collateral Demand,” *Journal of Financial Economics*, 2015, 116 (2), 237–256.
- European Central Bank, “Lessons learned from initial margin calls during the March 2020 market turmoil,” 2021.
- , “CCP initial margin models in Europe,” Occasional Paper Series 314, European Central Bank April 2023.
- European Securities and Markets Authority, “Final report on guidelines on CCP anti-procyclicality margin measures (ESMA70-151-1293),” Technical Report, ESMA 2018.
- , “Consultation paper on review of EMIR RTS with respect to procyclicality of CCP margin,” Technical Report, ESMA 2022.
- , “Final report: review of RTS No 153/2013 with respect to procyclicality of CCP margin,” Technical Report, ESMA 2023.
- European Systemic Risk Board, “Mitigating the procyclicality of margins and haircuts in derivatives markets and securities financing transactions,” Technical Report, ESRB January 2020.
- European Union, “Commission Delegated Regulation (EU) No 153/2013, Article 28 (Procyclicality),” 2013.
- Frazzini, Andrea, Ronen Israel, and Tobias J. Moskowitz, “Trading Costs,” Technical Report, AQR Capital Management 2018. Working paper, August 23, 2018.
- Gârleanu, Nicolae and Lasse Heje Pedersen, “Margin-Based Asset Pricing and Deviations from the Law of One Price,” *The Review of Financial Studies*, 2011, 24 (6), 1980–2022.
- Gatheral, Jim, “No-Dynamic-Arbitrage and Market Impact,” *Quantitative Finance*, 2010, 10 (7), 749–759.
- Geanakoplos, John, “The Leverage Cycle,” in Daron Acemoglu, Kenneth Rogoff, and Michael Woodford, eds., *NBER Macroeconomics Annual 2009, Volume 24*, University of Chicago Press, 2010, pp. 1–65.
- Glasserman, Paul and Qi Wu, “Persistence and Procyclicality in Margin Requirements,” *Management Science*, 2018, 64 (12), 5705–5724.
- Gorton, Gary and Andrew Metrick, “Securitized Banking and the Run on Repo,” *Journal of Financial Economics*, 2012, 104 (3), 425–451.

- Greenwood, Robin, Augustin Landier, and David Thesmar, "Vulnerable banks," *Journal of Financial Economics*, 2015, 115 (3), 471–485.
- Gromb, Denis and Dimitri Vayanos, "Equilibrium and Welfare in Markets with Financially Constrained Arbitrageurs," *Journal of Financial Economics*, 2002, 66 (2-3), 361–407.
- Gurrola-Perez, Pedro, "Procyclicality of CCP margin models: systemic problems need systemic approaches," WFE Working Paper, World Federation of Exchanges December 2020. Available at SSRN: 3779896.
- He, Zhiguo and Arvind Krishnamurthy, "Intermediary Asset Pricing," *American Economic Review*, 2013, 103 (2), 732–770.
- Hedegaard, Esben, "Causes and Consequences of Margin Levels in Futures Markets," Working Paper, Arizona State University 2014.
- Huang, Runhua and Elod Takáts, "Model risk at central counterparties: Is skin in the game a game changer?," *International Journal of Central Banking*, 2024, 20 (3).
- Huberman, Gur and Werner Stanzl, "Price Manipulation and Quasi-Arbitrage," *Econometrica*, 2004, 72 (4), 1247–1275.
- Jarrow, Robert A., "Market manipulation, bubbles, corners, and short squeezes," *Journal of Financial and Quantitative Analysis*, 1992, 27 (3), 311–336.
- Julliard, Christian, Gabor Pinter, Karamfil Todorov, Jean-Charles Wijnandts, and Kathy Yuan, "What Drives Repo Haircuts? Evidence from the UK Market," BIS Working Papers 1027, Bank for International Settlements November 2024. July 2022, revised November 2024.
- King, Thomas B., Travis D. Nesmith, Anna Paulson, and Todd Prono, "Central clearing and systemic liquidity risk," *International Journal of Central Banking*, 2023, 19 (4).
- Kiyotaki, Nobuhiro and John Moore, "Credit Cycles," *Journal of Political Economy*, 1997, 105 (2), 211–248.
- Kuong, John Chi-Fong and Vincent Maurin, "The design of a central counterparty," *Journal of Financial and Quantitative Analysis*, 2024, 59 (3), 1257–1299.
- Kyle, Albert S. and Anna Obizhaeva, "Market Microstructure Invariance: Empirical Hypotheses," *Econometrica*, 2016, 84 (4), 1345–1404.
- LCH, "LCH Ltd margin methodology," 2026.
- LCH Limited, "Procedures Section 2C: SwapClear Clearing Service," Rulebook / Procedures December 2025.
- LCH Ltd, "CPMI-IOSCO Self-Qualitative Assessment of Observance of the PFMI of LCH Ltd (Q3 2022)," 2022.

- LCH.Clearnet Limited, “4d(f) Request (CFTC) (describes PAIRS and volatility scaling via EWMA),” Technical Report, U.S. Commodity Futures Trading Commission November 2015.
- London Metal Exchange, “LME Clear SPAN Methodology (Version 1.1),” Technical Report, LME Clear August 2015.
- , “LME Clear margin parameter files,” 2026.
- , “Margin information,” 2026.
- Loon, Yee Cheng and Zhaodong (Ken) K. Zhong, “The impact of central clearing on counterparty risk, liquidity, and trading: Evidence from the credit default swap market,” *Journal of Financial Economics*, 2014, 112 (1), 91–115.
- Menkveld, Albert J. and Guillaume Vuillemeij, “The economics of central clearing,” *Annual Review of Financial Economics*, 2021, 13, 153–178.
- Murphy, David, Michalis Vasios, and Nicholas Vause, “An investigation into the procyclicality of risk-based initial margin models,” Financial Stability Paper 29, Bank of England 2014.
- , —, and —, “A comparative analysis of tools to limit the procyclicality of initial margin requirements,” Working Paper 597, Bank of England 2016.
- Nicolai, Francesco, “Dynamic Arbitrage from Price-Based Risk Constraints,” Technical Report, Working paper 2026. Working paper (February 2026).
- Onur, Esen and David Reiffen, “The Effect of Settlement Rules on the Incentive to Bang the Close,” Manuscript, U.S. Commodity Futures Trading Commission 2018.
- Pinter, Gabor, “An Anatomy of the 2022 Gilt Market Crisis,” Staff Working Paper 1019, Bank of England March 2023.
- Pirrong, Craig, “The Economics of Central Clearing: Theory and Practice,” ISDA Discussion Papers Series, International Swaps and Derivatives Association 2011.
- , “A Bill of Goods: CCPs and Systemic Risk,” *Journal of Financial Market Infrastructures*, 2014, 2 (3), 55–89.
- Reserve Bank of Australia, “Assessment of LCH Limited’s SwapClear Service: Appendix B (PAIRS description and procyclicality floor),” Technical Report, Reserve Bank of Australia 2019.
- Schneider, Michael and Fabrizio Lillo, “Cross-Impact and No-Dynamic-Arbitrage,” *Quantitative Finance*, 2019, 19 (1), 137–154.
- Shleifer, Andrei and Robert W. Vishny, “Fire Sales in Finance and Macroeconomics,” *Journal of Economic Perspectives*, 2011, 25 (1), 29–48.
- The Options Clearing Corporation, “Cross Margin Programs,” 2026.
- , “Margin Methodology,” 2026.

U.S. Commodity Futures Trading Commission, “CFTC Press Release 4542-01: Former Member of New York Futures Exchange Charged with Price Manipulation and False Representations,” Press Release July 2001.

—, “CFTC Press Release 4880-04: U.S. Commodity Futures Trading Commission Finds that Norman Eisler Manipulated Settlement Prices,” Press Release January 2004.

U.S. Commodity Futures Trading Commission, Division of Trading and Markets, “Report on Lessons Learned from the Failure of Klein & Co. Futures, Inc.,” Staff report July 2001.

U.S. Department of Justice, Office of the Solicitor General, “Brief for the United States as Amicus Curiae Supporting Petitioner, Klein & Co. Futures, Inc. v. Board of Trade of the City of New York, No. 06-1265,” Supreme Court amicus brief 2007.

U.S. Securities and Exchange Commission, “Release No. 34-100664: Notice of Filing of Proposed Rule Change by The Options Clearing Corporation,” August 2024.

—, “Release No. 34-99393: Notice of Filing of Proposed Rule Change by The Options Clearing Corporation,” January 2024.

Wang, Jessie Jiaxu, Agostino Capponi, and Hongzhong Zhang, “A Theory of Collateral Requirements for Central Counterparties,” *Management Science*, 2022, 68 (9), 6993–7017.

Internet Appendix
for
An Impossibility Theorem for Price-Based Risk Constraints

A Solving the general case

For simplicity and clarity, the main text specializes to realized variance and linear impact. The mechanism itself is local and does not require differentiability. It only requires that an admissible perturbation of the sampled transaction price at t that raises the risk input also raises posted margin by a nontrivial amount; that, in states where margins bind, this tightening induces a nontrivial adjustment in positions; and that the induced forced order flow moves transaction prices at first order. This Appendix states these local conditions without derivatives and then reruns the same two-step construction around a benchmark path with no deviation at time t . We write $(\cdot)^{(0)}$ for benchmark objects (prices, inputs, and margins) and compare them to outcomes under a small two-period round trip.

A.1 One-sided slopes

Definition 5 (One-sided strong increase). A function $f : \mathbb{R} \rightarrow \mathbb{R}$ is one-sided strongly increasing at x_0 with constant $c_f > 0$ if there exists $\delta_f > 0$ such that for all $h \in [0, \delta_f]$,

$$f(x_0 + h) - f(x_0) \geq c_f h. \quad (36)$$

Definition 6 (Directional sensitivity of the risk input in the last price). Let $\Gamma : \mathbb{R}^{m+1} \rightarrow \mathbb{R}$ and $\Gamma_t = \Gamma(P_{t-m}, \dots, P_t)$. Fix $(P_{t-m}, \dots, P_{t-1})$ and define $\Gamma^{\text{last}}(p) = \Gamma(P_{t-m}, \dots, P_{t-1}, p)$. We say that Γ is directionally sensitive at p_0 if there exist $s \in \{-1, +1\}$ and constants $c_\Gamma > 0$ and $\delta_\Gamma > 0$ such that for all $h \in [0, \delta_\Gamma]$,

$$\Gamma^{\text{last}}(p_0 + sh) - \Gamma^{\text{last}}(p_0) \geq c_\Gamma h. \quad (37)$$

Definition 5 replaces a right-derivative bounded away from zero with a one-sided linear lower bound. Definition 6 is the only step where the realized-variance structure in the main text mattered: it ensures that an admissible perturbation of the last sampled mark can raise the risk input at first order; conditional on that, the remainder of the construction is unchanged.

A.2 Local assumptions

Fix a reachable benchmark configuration at time t and interpret all statements below locally at that benchmark. Let $P_t^{(0)}$ denote the benchmark transaction price at t absent a deviation, and let $\Gamma_t^{(0)}$ denote the corresponding risk input.

Assumption 11 (Risk input is directionally sensitive). The risk input Γ is directionally sensitive at the benchmark last mark $P_t^{(0)}$ in the sense of Definition 6. Moreover, $\Gamma_t^{(0)}$ lies in the interior of an interval I_Γ on which the posted margin rule is evaluated.

Assumption 12 (Posted margin has a one-sided slope bound). The posted margin rule is $M_{t+1} = g(\Gamma_t)$, where $g : \mathbb{R} \rightarrow (0, \infty)$ is weakly increasing. There exists $c_g > 0$ such that for all $x, y \in I_\Gamma$ with $y \geq x$,

$$g(y) - g(x) \geq c_g(y - x). \quad (38)$$

Assumption 13 (Forced liquidation has a local lower bound). There exist constants $c_X > 0$ and $\delta_M > 0$ such that for any margin increase $\Delta M \in [0, \delta_M]$, the induced reduction in aggregate constrained position satisfies

$$L(\Delta M) = X(M_t) - X(M_t + \Delta M) \geq c_X \Delta M. \quad (39)$$

Assumption 14 (Generic price impact). Transaction prices satisfy

$$P_u = V_u + \Psi_u(Q_u), \quad Q_u = \Delta X_u + a_u, \quad (40)$$

where $\Psi_u(0) = 0$ and Ψ_u is odd and weakly increasing. There exist constants $0 < \underline{\alpha}_u \leq \bar{\alpha}_u < \infty$ and $\delta_\Psi > 0$ such that for all $q \in [0, \delta_\Psi]$,

$$\underline{\alpha}_u q \leq \Psi_u(q) \leq \bar{\alpha}_u q. \quad (41)$$

By oddness, (41) also implies $-\bar{\alpha}_u |q| \leq \Psi_u(q) \leq -\underline{\alpha}_u |q|$ for $q \in [-\delta_\Psi, 0]$. Assumption 14 contains linear impact as a special case and covers any monotone impact schedule with local slopes bounded away from zero and infinity on the relevant flow range.

A.2.1 Broader constrained demand schedules and margin elasticity

This subsection gives a sufficient condition for Assumption 13. The main text uses a target-with-cap rule to make margin elasticity explicit; the same local slope obtains for a broad class of truncated demands whenever the margin cap binds for a positive mass of traders with nontrivial total equity.

Fix time t and a benchmark state. Let $\tilde{x}_t(i)$ denote trader i 's unconstrained desired position and let $E(i)$ denote equity. The implemented position under margin level $M > 0$ is

$$x_t(i; M) = \text{sgn}(\tilde{x}_t(i)) \min \left\{ |\tilde{x}_t(i)|, \frac{E(i)}{M} \right\}. \quad (42)$$

Aggregate constrained demand is $X(M) = \int_0^1 x_t(i; M) di$.

Lemma 11 (Sufficient condition for Assumption 13). Fix a benchmark margin $M_t > 0$ and let $B \subseteq [0, 1]$ be measurable such that

$$|\tilde{x}_t(i)| > \frac{E(i)}{M_t} \text{ for all } i \in B, \quad \int_B E(i) di > 0, \quad (43)$$

and $\text{sgn}(\tilde{x}_t(i))$ is constant on B . Then for any $\delta_M > 0$ and all $\Delta M \in [0, \delta_M]$,

$$X(M_t) - X(M_t + \Delta M) \geq \left(\frac{\int_B E(i) di}{(M_t + \delta_M)^2} \right) \Delta M. \quad (44)$$

In particular, Assumption 13 holds with $c_X = \frac{\int_B E(i) di}{(M_t + \delta_M)^2}$.

Proof. Fix $\Delta M \in [0, \delta_M]$ and write $M' = M_t + \Delta M$. For $i \in B$, the cap binds at M_t by (43), and since $M' \geq M_t$ it also binds at M' . Hence

$$x_t(i; M_t) = \text{sgn}(\tilde{x}_t(i)) \frac{E(i)}{M_t}, \quad x_t(i; M') = \text{sgn}(\tilde{x}_t(i)) \frac{E(i)}{M'}.$$

With constant sign on B ,

$$x_t(i; M_t) - x_t(i; M') = \frac{E(i)\Delta M}{M_t M'}.$$

Integrating over B and using $M_t M' \leq (M_t + \delta_M)^2$ yields (44). \square

Demand (42) arises from any concave problem with a unique unconstrained maximizer $\tilde{x}_t(i)$ and a margin constraint $|x|M \leq E(i)$. Lemma 11 shows that if the cap binds for a positive-mass set with nontrivial total equity, then aggregate positions respond to margin with a strictly positive local slope. The bound increases with $\int_B E(i) di$ and decreases with the relevant margin level M_t .

A.2.2 A two-step round trip in the general case

To keep the profit bound clean under a fully nonparametric Ψ_{t+1} , we impose the same local benchmark configuration as in the main text.

Assumption 15. At the benchmark S_t : (i) $\Delta X_t = 0$; (ii) letting ΔX_{t+1}^0 denote constrained order flow at $t+1$ absent a deviation at t , we have $\Delta X_{t+1}^0 = 0$; (iii) $\mathbb{E}[V_{t+1} \mid S_t] = V_t$.

Lemma 12 (A profitable two-period round trip under non-smooth primitives). *Suppose Assumptions 11–15 hold at a reachable benchmark state S_t . For concreteness, assume the risk input increases when the last sampled mark is pushed down (in Definition 6, $s = -1$).¹⁷*

Let $\bar{q}_t > 0$ be such that for every $q \in (0, \bar{q}_t]$, the round trip $(a_t, a_{t+1}) = (-q, q)$ keeps all local restrictions in force: (i) $\Psi_t(q) \leq \delta_\Gamma$; (ii) the perturbed input remains in I_Γ ; (iii) the induced margin change $\Delta M_{t+1} = g(\Gamma_t) - g(\Gamma_t^{(0)})$ lies in $[0, \delta_M]$; (iv) the induced liquidation $L = X(M_t) - X(M_t + \Delta M_{t+1})$ satisfies $L - q \in [0, \delta_\Psi]$; and (v) when imposed, the hard funding constraint (12) holds (equivalently, $q \leq \bar{F}/M_t$). Define

$$q_{\max, t} := \min\{\bar{a}, \delta_\Psi, \bar{q}_t\}.$$

Fix any $q \in (0, q_{\max, t}]$ and consider $a_t = -q$, $a_{t+1} = q$. Then

$$\Delta \Gamma_t := \Gamma_t - \Gamma_t^{(0)} \geq c_\Gamma \underline{\alpha}_t q, \quad \Delta M_{t+1} \geq c_g c_\Gamma \underline{\alpha}_t q, \quad L \geq c_X c_g c_\Gamma \underline{\alpha}_t q.$$

If, in addition, the gain condition

$$\underline{\alpha}_{t+1} c_X c_g c_\Gamma \underline{\alpha}_t > \bar{\alpha}_t + \underline{\alpha}_{t+1} \tag{45}$$

holds, then the conditional expected profit satisfies

$$\mathbb{E}[\Pi_1 \mid S_t] \geq \left(\underline{\alpha}_{t+1} c_X c_g c_\Gamma \underline{\alpha}_t - (\bar{\alpha}_t + \underline{\alpha}_{t+1}) \right) q^2 - 2\tau q. \tag{46}$$

Let

$$A = \underline{\alpha}_{t+1} c_X c_g c_\Gamma \underline{\alpha}_t - (\bar{\alpha}_t + \underline{\alpha}_{t+1}), \quad q_{\min} = \frac{2\tau}{A} \text{ when } A > 0.$$

Under (45), $A > 0$, and if $q_{\max, t} > q_{\min}$ there exists an admissible $q \in (q_{\min}, q_{\max, t}]$ such that $\mathbb{E}[\Pi_1 \mid S_t] > 0$.

¹⁷If $s = +1$, reverse the signs of the round trip and use the buy-in analogue of Assumption 13; see the sign-reversal discussion following Theorem 1.

Proof. Under Assumption 15(i), $Q_t = a_t = -q$, so

$$P_t = V_t + \Psi_t(-q) = V_t - \Psi_t(q), \quad P_t^{(0)} = V_t,$$

and by (41), $\Psi_t(q) \geq \underline{\alpha}_t q$. With $s = -1$, Assumption 11 yields $\Delta\Gamma_t \geq c_\Gamma \Psi_t(q) \geq c_\Gamma \underline{\alpha}_t q$. Since $\Gamma_t, \Gamma_t^{(0)} \in I_\Gamma$ and $\Gamma_t \geq \Gamma_t^{(0)}$, Assumption 12 implies $\Delta M_{t+1} \geq c_g \Delta\Gamma_t \geq c_g c_\Gamma \underline{\alpha}_t q$, and then Assumption 13 gives $L \geq c_X \Delta M_{t+1} \geq c_X c_g c_\Gamma \underline{\alpha}_t q$. Under Assumption 15(ii), the constrained sector has zero net flow at $t+1$ absent the deviation; with the deviation it sells L while the deviator buys q , so $Q_{t+1} = q - L$ and $P_{t+1} = V_{t+1} + \Psi_{t+1}(q - L)$. Condition (45) implies $c_X c_g c_\Gamma \underline{\alpha}_t > 1$, hence $L > q$ and $q - L < 0$. By the definition of \bar{q}_t , $L - q \in [0, \delta_\Psi]$, so oddness and (41) give

$$P_{t+1} = V_{t+1} - \Psi_{t+1}(L - q) \leq V_{t+1} - \underline{\alpha}_{t+1}(L - q), \quad P_t \geq V_t - \bar{\alpha}_t q.$$

Therefore

$$P_t - P_{t+1} \geq (V_t - V_{t+1}) - \bar{\alpha}_t q + \underline{\alpha}_{t+1}(L - q).$$

Under the round trip, $\Pi_1 = q(P_t - P_{t+1}) - 2\tau q$, so taking conditional expectations and using Assumption 15(iii) yields

$$\mathbb{E}[\Pi_1 \mid S_t] \geq q \left(-\bar{\alpha}_t q + \underline{\alpha}_{t+1}(L - q) \right) - 2\tau q.$$

Substituting $L \geq c_X c_g c_\Gamma \underline{\alpha}_t q$ gives (46). Under (45), $A > 0$ and the lower bound equals $Aq^2 - 2\tau q = q(Aq - 2\tau)$, which is positive whenever $q > 2\tau/A$. If $q_{\max,t} > q_{\min} = 2\tau/A$, pick any $q \in (q_{\min}, q_{\max,t}]$. \square

Assumption 15(ii) is a convenience: it avoids imposing structure on Ψ_{t+1} beyond local slope bounds. With mild local regularity of Ψ_{t+1} on the support of baseline flow (for example local affinity on that range), a weaker mean-zero condition such as $\mathbb{E}[\Delta X_{t+1}^0 \mid S_t] = 0$ suffices, as in the main text.

A.2.3 The general impossibility statement

Lemma 12 gives a lower bound on the conditional expected profit of a two-period trigger-and-reverse deviation under local one-sided conditions. If the gain condition (45) holds and the admissible local range contains some q above the execution-cost threshold q_{\min} , the bound is strictly positive and yields a profitable round trip.

Theorem 5 (General impossibility theorem). *Fix a Markov equilibrium with price formation (40) and posted margin rule $M_{t+1} = g(\Gamma_t)$. Suppose there exists a reachable benchmark state S_t at which liquidity continuity holds (Definition 4) and Assumptions 11–15 hold. Let $q_{\max,t}$ and q_{\min} be defined as in Lemma 12. If (45) holds and $q_{\max,t} > q_{\min}$, then the mechanism is not round-trip manipulation-proof at S_t in the sense of Definition 3 with $\rho = \kappa = 0$: there exists an admissible two-period round trip with $\mathbb{E}[\Pi_1 \mid S_t] > 0$.*

Consequently, at S_t the following cannot all hold simultaneously: one-sided risk sensitivity in the sense of Assumption 12, liquidity continuity (Definition 4), and round-trip manipulation-proofness net of execution costs (Definition 3 with $\rho = \kappa = 0$). Equivalently, at any reachable S_t with liquidity continuity, round-trip manipulation-proofness implies that at least one of Assumptions 11–15, the gain inequality (45), or the feasibility condition $q_{\max,t} > q_{\min}$ fails.

Proof. Under the stated hypotheses, Lemma 12 implies that any $q \in (q_{\min}, q_{\max, t}]$ yields an admissible two-period round trip $(a_t, a_{t+1}) = (-q, q)$ with $\mathbb{E}[\Pi_1 \mid S_t] > 0$, contradicting round-trip manipulation-proofness at S_t (net of execution costs). \square

A.3 Verifying directional sensitivity for standard CCP-style inputs

Assumption 11 is intentionally weak. We do not model full CCP implementation details; we only give simple sufficient conditions ensuring that a risk input computed from recent marks is directionally sensitive in the last sampled mark.

A.3.1 Smooth functionals of recent returns

Let $R_j = P_j - P_{j-1}$ and suppose the input is a smooth function of the most recent m returns,

$$\Gamma_t = \phi(R_{t-m+1}, \dots, R_t),$$

for a continuously differentiable map ϕ defined on an open neighborhood of the realized return vector.

Lemma 13. *If ϕ is continuously differentiable near the realized return vector and $\partial\phi/\partial R_t \neq 0$ at the realized point, then Γ is directionally sensitive at the benchmark last mark in the sense of Definition 6.*

Proof. Hold $(P_{t-m}, \dots, P_{t-1})$ fixed and perturb only the last mark: $P_t = p_0 \mapsto p_0 + h$. Then only the last return changes, $R_t \mapsto R_t + h$. Define $u(h) := \phi(R_{t-m+1}, \dots, R_{t-1}, R_t + h)$, so $\Gamma^{\text{last}}(p_0 + h) = u(h)$. By differentiability, $u'(0) = \partial\phi/\partial R_t \neq 0$. Let $s = \text{sgn}(u'(0))$. Continuity of u' implies that for some $\delta > 0$, $u'(sh)s \geq |u'(0)|/2$ for all $h \in [0, \delta]$. The mean value theorem then gives $u(sh) - u(0) = u'(\xi)sh \geq (|u'(0)|/2)h$ for some $\xi \in (0, sh)$. This is Definition 6 with $c_\Gamma = |u'(0)|/2$ and $\delta_\Gamma = \delta$. \square

Realized variance is the special case $\phi(r) = \sum_j r_j^2$.

A.3.2 Parametric VaR/ES built from volatility estimates

A common parametric form is $\Gamma_t = k_p \hat{\sigma}_t$ where $\hat{\sigma}_t = h(R_{t-m+1}, \dots, R_t)$ and $k_p > 0$ is a fixed multiplier.

Lemma 14. *If h is continuously differentiable near the realized return vector and $\partial h/\partial R_t \neq 0$ at that point, then $\Gamma_t = k_p \hat{\sigma}_t$ is directionally sensitive at the benchmark last mark in the sense of Definition 6.*

Proof. Apply Lemma 13 with $\phi = k_p h$. Since $k_p > 0$, $\partial\phi/\partial R_t = k_p(\partial h/\partial R_t) \neq 0$, so directional sensitivity follows. \square

A.4 Transient impact and longer horizons

The two-period round trip used in the main argument is the shortest horizon that can (i) perturb the sampled marks that enter the next update and then (ii) trade against the forced liquidation induced by the tighter requirement. Two extensions clarify what the proof does and does not rely on.

First, the mechanism does not require permanent impact. Allow P_{t+1} to depend on the history of total order flow (Q_{t+1}, Q_t, \dots) , but assume that the contemporaneous mapping $Q_{t+1} \mapsto P_{t+1}$ has a strictly positive local one-sided slope around the realized flow at $t+1$, in the sense of Assumption 14. Then the proof of Lemma 12 goes through after replacing $\underline{\alpha}_{t+1}$ by this instantaneous local slope. Additional dependence of P_{t+1} on earlier flows enters as extra terms; these do not remove the rule-induced predictable component created by forced liquidation, and any persistence of the initial leg typically makes the reversal leg cheaper conditional on the same liquidation response. The argument therefore needs only that liquidation flow moves the transaction price at first order when it arrives.

Second, longer-horizon round trips are not needed for the impossibility result, but they can strengthen it. If a profitable two-period round trip exists at a reachable state, it can be embedded into any longer horizon by setting $a_{t+2} = \dots = a_{t+T} = 0$, which preserves admissibility and preserves the same expected profit. Longer horizons can also generate additional profitable deviations when successive updates propagate the initial perturbation forward (for example, because prices affected at $t+1$ enter subsequent inputs and the relevant one-sided slopes remain active). A full finite-horizon characterization of such multi-update trigger-and-reverse strategies, and the corresponding stronger stress tests for a fixed disclosed rulebook, is developed in Nicolai (2026). For the purposes of this paper, a single profitable two-period deviation is sufficient to violate manipulation-proofness at the evaluated state, so the slope cap used in Section 6 and Appendix C should be read as a tractable sufficient restriction based on the minimal deviation. Enforcing the full finite-horizon requirement in Definition 3 against multi-update strategies can only tighten the admissible local pass-through bound, and thus weakly expands the pooling regions in the optimal schedule.

A.5 Linear execution and funding costs

Let $\tau \geq 0$ denote proportional execution costs, so each trade a_t incurs cost $\tau|a_t|$. Let $\rho \geq 0$ denote a linear funding wedge, so carrying inventory y_t over $(t, t+1]$ incurs cost $\rho M_t|y_t|$. For a trade sequence $a_{0:T}$, define

$$\Pi_T^{\tau, \rho, \kappa}(a_{0:T}) = - \sum_{t=0}^T a_t P_t - \tau \sum_{t=0}^T |a_t| - \rho \sum_{t=0}^{T-1} M_t |y_t| - \kappa \sum_{t=0}^{T-1} |y_t|^2.$$

When $\tau = \rho = 0$, this reduces to Π_T^κ in (11). The optional hard funding constraint (12) continues to impose an upper bound on admissible trigger sizes.

Lemma 15. *Fix a reachable state S_t satisfying Assumptions 5, 1, 3, and 4. Suppose the maximum admissible trigger size is $q_{\max} > 0$, i.e., every $q \in (0, q_{\max}]$ satisfies the locality restrictions (and, if imposed, the hard funding constraint). Let*

$$\underline{G} = \alpha c_X c_g c_\Gamma,$$

where c_Γ is the local sensitivity constant in Assumption 5. For the two-period round trip $a_t = -q, a_{t+1} = q$,

$$\mathbb{E}[\Pi_1^{\tau, \rho, \kappa} \mid S_t] \geq Aq^2 - Bq, \quad A = \alpha(\underline{G} - 2) - \kappa, \quad B = 2\tau + \rho M_t. \quad (47)$$

If $A > 0$ and $q_{\min} := B/A$, then $\mathbb{E}[\Pi_1^{\tau, \rho, \kappa} \mid S_t] > 0$ for every admissible $q \in (q_{\min}, q_{\max}]$. Thus linear wedges add the feasibility condition $q_{\max} > q_{\min}$.

Proof. As in Lemma 5, the induced forced liquidation satisfies $L \geq c_X c_g c_\Gamma \alpha q$. Under $\mathbb{E}[V_{t+1} \mid S_t] = V_t$ and $\mathbb{E}[\Delta X_{t+1}^0 \mid S_t] = 0$,

$$\mathbb{E}[\Pi_1^{\tau, \rho, \kappa} \mid S_t] = \alpha q(L - 2q) - (2\tau + \rho M_t)q - \kappa q^2.$$

Substituting the lower bound for L yields (47). If $A > 0$, then $Aq^2 - Bq > 0$ is equivalent to $q > q_{\min} = B/A$. \square

B Strategic constrained traders and endogenous participation

Sections 2.4–2.5 treat the constrained sector’s target exposure \bar{x}_t as fixed at the start of the pricing window. This reduced form makes the forced-liquidation channel transparent, but it raises a natural question: if constrained traders anticipate that within-window trading can move Γ_t and hence $M_{t+1} = g(\Gamma_t)$, will they scale back ex ante and eliminate amplification?

We answer in two steps. First, we microfound target choice when traders take g and the manipulability of Γ_t as given: optimal targets fall with the posted margin and with the anticipated dispersion of manipulation-driven margin changes, reducing the mass of agents near the binding cap and lowering the effective local liquidation elasticity. Second, we show that this is dampening, not a fix. The main theorem applies with c_X replaced by an equilibrium object c_X^{eff} that aggregates any anticipatory repositioning or within-window re-optimization. The mechanism disappears only in degenerate regimes in which constrained participation collapses or the requirement is kept slack in the relevant states, which comes at a market-functioning and coverage cost.

B.1 A two-period game

Fix time t and consider the two-period horizon $(t, t + 1)$ in the environment of Section 2. The CCP commits to the posted rule $M_{t+1} = g(\Gamma_t)$, where Γ_t is computed from sampled transaction prices inside period t (Section 2.3). The sampling protocol is common knowledge and, because it depends on tradable marks, can be influenced by trades within the window.

1. At the start of t , the public state S_t is observed (including M_t and the lagged objects entering Γ_t). Each constrained trader $i \in [0, 1]$ chooses a target exposure $\bar{x}_t(i)$, and the implemented position is

$$x_t(i) = \pi(\bar{x}_t(i), E(i)/M_t) = \text{sgn}(\bar{x}_t(i)) \min \left\{ |\bar{x}_t(i)|, \frac{E(i)}{M_t} \right\},$$

as in (6). The difference from the main text is that $\bar{x}_t(i)$ is chosen optimally as a function of S_t and the known rule g .

2. The strategic trader chooses a_t inside the pricing window. Prices satisfy (2), so a_t affects the sampled marks and hence Γ_t .
3. The CCP posts $M_{t+1} = g(\Gamma_t)$. Constrained traders adjust at $t + 1$ to satisfy the margin constraint under M_{t+1} ; when $M_{t+1} > M_t$ and the constraint binds, this adjustment is forced liquidation and generates price-impactful order flow (Section 2.2).
4. The strategic trader chooses a_{t+1} and completes the round trip, $a_t + a_{t+1} = 0$.

B.2 Objective

We use a parsimonious objective that captures three elements: a genuine exposure motive, a convex penalty for large positions, and costs from posted margin, including both funding costs and anticipated losses from manipulation-induced margin changes.

For transparency, focus on the long side in a neighborhood where forced adjustment is selling.¹⁸ Fix a state S_t . Constrained trader i chooses $\bar{x}_t(i) \geq 0$ to maximize

$$\max_{\bar{x} \geq 0} \left\{ b_t(i) x(\bar{x}) - \frac{\chi}{2} x(\bar{x})^2 - \rho M_t x(\bar{x}) - \frac{\eta}{2} \sigma_{M,t}^2(S_t) x(\bar{x})^2 \right\}, \quad (48)$$

where $x(\bar{x}) = \min\{\bar{x}, E(i)/M_t\}$ is the end-of- t position, $b_t(i) \geq 0$ is the marginal value of exposure (hedging or expected return), $\chi > 0$ is a baseline quadratic penalty, and $\rho \geq 0$ is the per-period funding spread for posted margin (Appendix A.5). The final term is a reduced-form expected loss from exposure to margin jumps: when Γ_t is moved within the window, M_{t+1} shifts and binding constraints force adjustment, generating predictable price pressure at $t + 1$; larger positions scale these losses, motivating the quadratic form. The key equilibrium object is

$$\sigma_{M,t}^2(S_t) = \mathbb{E}[(M_{t+1} - M_t)^2 \mid S_t], \quad (49)$$

the conditional second moment of the next margin change under equilibrium behavior.

To close the manipulator side explicitly, suppose the strategic trader chooses the trigger size $q \geq 0$ in the two-leg round trip $(a_t, a_{t+1}) = (-q, q)$ to maximize conditional expected funding-adjusted profit (as in (11), net of proportional execution costs and any modeled funding or quadratic inventory costs), subject to admissibility. Given S_t and the constrained-sector targets chosen at step 1, the trigger trade moves the last sampled mark through impact, which shifts the sampled input and therefore the posted margin. A local chain-rule approximation is

$$\Delta M_{t+1} \approx g'(\Gamma_t) \cdot \frac{\partial \Gamma_t}{\partial P_t} \cdot \frac{\partial P_t}{\partial q} \cdot q, \quad \text{with } \frac{\partial P_t}{\partial q} = -\alpha \text{ under (2) for } a_t = -q.$$

If the strategic trader plays a pure best response $q^*(S_t)$, then conditional on S_t the next margin is pinned down by $M_{t+1} = g(\Gamma_t)$ evaluated at $q^*(S_t)$, so $\sigma_{M,t}^2(S_t) = (M_{t+1} - M_t)^2$; if he mixes over q (or over the trigger sign), then $\sigma_{M,t}^2(S_t)$ is the corresponding conditional second moment. In either case, $\sigma_{M,t}^2(S_t)$ is endogenous through the manipulator's equilibrium behavior.

Because $M_{t+1} = g(\Gamma_t)$ and Γ_t is computed from sampled transaction prices, $\sigma_{M,t}^2(S_t)$ rises with local pass-through in g and with the susceptibility of the sampling window to within-window trading. This increases the constrained sector's ex ante cost of holding large positions in states where constraints may bind. If the margin cap does not bind for trader i , so $x(\bar{x}) = \bar{x}$, then (48) has the interior solution

$$\bar{x}_t^*(i) = \frac{(b_t(i) - \rho M_t)_+}{\chi + \eta \sigma_{M,t}^2(S_t)}. \quad (50)$$

Three comparative statics follow. First, $\bar{x}_t^*(i)$ falls with M_t whenever $b_t(i) > \rho M_t$, with

$$\frac{\partial \bar{x}_t^*(i)}{\partial M_t} = -\frac{\rho}{\chi + \eta \sigma_{M,t}^2(S_t)} < 0.$$

Second, holding M_t fixed, $\bar{x}_t^*(i)$ falls with $\sigma_{M,t}^2(S_t)$. Third, if M_t is high or $\sigma_{M,t}^2(S_t)$ is large enough that $b_t(i) \leq \rho M_t$, then $\bar{x}_t^*(i) = 0$: the trader exits. When the cap binds, the realized position is pinned

¹⁸With sign reversals, the same analysis applies to forced buy-ins when short constraints tighten.

at $x_t(i) = E(i)/M_t$ and the shadow cost of binding rises with $\sigma_{M,t}^2(S_t)$, so higher manipulation risk reduces participation and shrinks the mass of agents near the cap.

B.3 Aggregate demand and an effective liquidation elasticity

Let $\bar{x}_t^*(i)$ denote the optimal target in (50) and define the realized end-of- t position

$$x_t^*(i) = \min \left\{ \bar{x}_t^*(i), \frac{E(i)}{M_t} \right\}.$$

Aggregate constrained exposure is

$$X_t^* = \int_0^1 x_t^*(i) di.$$

Define the effective local liquidation elasticity at the state S_t by

$$c_X^{\text{eff}}(S_t) := -\frac{\partial}{\partial M} X(M, \bar{x}_t^*) \Big|_{M=M_t}, \quad (51)$$

where $X(M, \bar{x})$ is the aggregate mapping in (7) and the derivative is taken holding the equilibrium targets $\{\bar{x}_t^*(i)\}_{i \in [0,1]}$ fixed at their start-of- t values. This is the object that enters the two-period round-trip computation: within the pricing window in period t , the strategic trader can move Γ_t and hence M_{t+1} , but the constrained sector's targets are already chosen.

If constrained traders can adjust within the pricing window after observing contemporaneous prices or manipulation signals, the relevant object is the derivative of the aggregate within-window best response with respect to the next posted margin M_{t+1} . Under standard concavity of the constrained trader objective, this best-response slope is weakly smaller in absolute value than the fixed-target slope in (51), and it equals zero only if the constrained set is locally empty (slack constraints or exit). Hence allowing within-window adjustment can only reduce c_X^{eff} , and the round-trip channel persists whenever the equilibrium constrained mass is positive. Let

$$m_t(S_t) = \int_0^1 \mathbf{1} \left\{ \bar{x}_t^*(i) \geq \frac{E(i)}{M_t} \right\} di$$

denote the mass of traders whose cap binds at (S_t, M_t) . If $m_t(S_t) > 0$, then differentiating (7) with respect to M yields

$$c_X^{\text{eff}}(S_t) = \frac{1}{M_t^2} \int_0^1 E(i) \mathbf{1} \left\{ \bar{x}_t^*(i) \geq \frac{E(i)}{M_t} \right\} di > 0. \quad (52)$$

Equation (50) implies that $\bar{x}_t^*(i)$ is decreasing in $\sigma_{M,t}^2(S_t)$, so higher anticipated manipulation risk shrinks the indicator set in (52) and therefore lowers $c_X^{\text{eff}}(S_t)$. This is endogenous dampening: anticipation reduces the mass of constrained agents close to the binding region and weakens the liquidation response induced by a given margin shock.

B.4 Anticipation dampens but does not eliminate the channel

We now link the strategic constrained sector to the impossibility logic.

Lemma 16. Fix a state S_t and consider the two-period game above. If $m_t(S_t) > 0$ (equivalently, $c_X^{\text{eff}}(S_t) > 0$), then there exist $\delta_X > 0$ and $c_X^{\text{eff}} > 0$ such that, holding the equilibrium targets fixed, the aggregate mapping $M \mapsto X(M, \bar{x}_t^*)$ is continuously differentiable on $[M_t, M_t + \delta_X]$ and satisfies

$$\frac{\partial X}{\partial M}(M, \bar{x}_t^*) \leq -c_X^{\text{eff}} \quad \text{for all } M \in [M_t, M_t + \delta_X].$$

In particular, Assumption 3 holds with c_X replaced by c_X^{eff} .

Proof. For long positions, $X(M, \bar{x}_t^*) = \int_0^1 \min\{\bar{x}_t^*(i), E(i)/M\} di$ is continuously differentiable in M on any interval over which the binding set does not change. Since $m_t(S_t) > 0$ at $M = M_t$, there exists $\delta_X > 0$ such that the binding set remains non-empty on $[M_t, M_t + \delta_X]$. Differentiating under the integral yields

$$\frac{\partial X}{\partial M}(M, \bar{x}_t^*) = -\frac{1}{M^2} \int_0^1 E(i) \mathbf{1}\left\{\bar{x}_t^*(i) \geq \frac{E(i)}{M}\right\} di.$$

The indicator set has positive mass throughout the neighborhood and $E(i) > 0$, so the right-hand side is bounded above by $-c_X^{\text{eff}}$ for some $c_X^{\text{eff}} > 0$ on $[M_t, M_t + \delta_X]$. \square

Lemma 16 shows that endogenizing targets preserves the key local ingredient: whenever some traders are locally margin-binding, margin increases still generate first-order forced liquidation, with slope summarized by c_X^{eff} .

Proposition 3 (Anticipation dampens but does not eliminate the impossibility logic). Fix a margin schedule g and a state S_t at which the local risk-input sensitivity and slope conditions of the main theorem hold, and consider the two-period game above. If $m_t(S_t) > 0$ (so $c_X^{\text{eff}}(S_t) > 0$), then the two-period deviation in Lemma 5 yields strictly positive conditional expected funding-adjusted profit whenever the amplification condition holds with c_X replaced by $c_X^{\text{eff}}(S_t)$. Thus anticipation can attenuate amplification by lowering c_X^{eff} , but it cannot restore joint feasibility of risk sensitivity, liquidity continuity, and manipulation-proofness in any state where constraints still bind locally. If instead $m_t(S_t) = 0$ (equivalently, $c_X^{\text{eff}}(S_t) = 0$), then forced liquidation is locally absent and the round-trip channel is shut down; avoidance occurs only through slack constraints or participation collapse in those states.

Proof. Lemma 5 uses two ingredients: (i) the deviation shifts the input and hence M_{t+1} at first order, and (ii) a marginal increase in M_{t+1} induces first-order forced liquidation and hence a predictable price component at $t + 1$ through (2). Lemma 16 provides (ii) with c_X^{eff} in place of c_X . Substituting into the algebra in Lemma 5 yields the same profit bound with c_X^{eff} replacing c_X . If $m_t(S_t) = 0$, then $c_X^{\text{eff}} = 0$ and the forced-liquidation leg is locally absent. \square

B.5 Design implications and the participation trade-off

The strategic extension highlights an additional cost of steep pass-through in g . A larger local slope increases the direct profitability of within-window manipulation, as in the main theorem, but it also increases the constrained sector's perceived exposure to manipulation-driven margin jumps, raising $\sigma_{M,t}^2(S_t)$ and lowering equilibrium targets in (50). The resulting decline in c_X^{eff} is stabilizing, but it

operates through reduced leveraged participation and thus through weaker liquidity provision and less effective use of the cleared market. This reinforces the design logic in Section 6. Slope caps and pooling regions reduce manipulation incentives and, at the same time, attenuate ex ante contractions in constrained participation. Pooling-plus-adjustment schedules can therefore be read not only as a tractable way to satisfy the incentive constraint emphasized in the main theorem, but also as a way to preserve participation while limiting amplification in binding, risk-sensitive updates.

C Optimal Margin Design

This appendix formalizes the single-asset margin-design problem sketched in Section 6. A CCP (or any margin-setting institution) posts next-period initial margin as a deterministic schedule

$$M_{t+1} = g(\Gamma_t),$$

where Γ_t is a price-based risk input computed from recent marks. The main theorem highlights two mechanical facts: because Γ_t is built from tradable marks, it can be moved at first order within the sampling window; and when the requirement binds for a nontrivial constrained sector, marginal increases in M_{t+1} induce forced deleveraging that feeds back into prices through impact. In binding states, the mapping from Γ_t to M_{t+1} is therefore a market-design object: it must trade off loss coverage, the market-functioning costs of margin updates, and the level cost of collateral, subject to an implementability restriction that limits short-horizon pass-through. The portfolio-margining extension with cross-impact begins in Section C.3 and is unchanged.

C.1 Losses, scale model, and stationary reduction

Let $D_{t+1} \geq 0$ denote next-period liquidation/default loss over $(t, t+1]$ that M_{t+1} is intended to cover. We connect losses to the risk input through a square-root scale model.

Assumption 16 (Square-root scale model). Along the benchmark path, there exists $Y_{t+1} \geq 0$ such that

$$D_{t+1} = \sqrt{\Gamma_t} Y_{t+1}, \quad (53)$$

and the conditional law of Y_{t+1} given the public state is independent of the level of Γ_t .

Assumption 16 implies quantile scaling: for any $p \in (0, 1)$,

$$Q_p(D_{t+1} \mid \Gamma_t = \gamma) = \sqrt{\gamma} Q_p(Y_{t+1}). \quad (54)$$

The CCP chooses a schedule $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ and posts $M_{t+1} = g(\Gamma_t)$. We evaluate performance under the benchmark stationary law of $(\Gamma_{t-1}, \Gamma_t, D_{t+1})$ and suppress $(\cdot)^{(0)}$. The per-period loss is

$$L_t(g) = \ell(D_{t+1} - g(\Gamma_t)) + \lambda \mathcal{C}(-c_X(g(\Gamma_t) - g(\Gamma_{t-1}))) + \kappa_M g(\Gamma_t), \quad (55)$$

where ℓ penalizes shortfall (coverage), \mathcal{C} penalizes forced-flow impulses from margin updates (market functioning), $c_X > 0$ is the local margin elasticity, and $\kappa_M > 0$ is the level cost of posted margin.¹⁹

With discount factor $\beta \in (0, 1)$, the infinite-horizon problem reduces under stationarity to a one-period criterion:

$$\min_{g \in \mathcal{G}} J(g), \quad J(g) := \mathbb{E}[L_t(g)]. \quad (56)$$

To obtain a tractable continuous program, specialize $\mathcal{C}(z) = |z|$ and linearize one-step updates for

¹⁹The normalization measures κ_M in units of marginal shortfall loss. For $\ell(x) = (x)_+$, this places κ_M in $[0, 1]$ when the minimizer is interior.

absolutely continuous g :

$$\lambda \mathcal{C}(-c_X(g(\Gamma_t) - g(\Gamma_{t-1}))) \approx \lambda c_X g'(\Gamma_{t-1}) |\Delta\Gamma_t|, \quad \Delta\Gamma_t := \Gamma_t - \Gamma_{t-1}. \quad (57)$$

Define the expected absolute one-step change

$$\delta(\gamma) = \mathbb{E}[|\Gamma_{t+1} - \Gamma_t| | \Gamma_t = \gamma]. \quad (58)$$

Replacing $|\Delta\Gamma_t|$ by its conditional mean yields the approximation

$$\min_{g \in \tilde{\mathcal{G}}} \tilde{J}(g), \quad \tilde{J}(g) := \mathbb{E}[\ell(D_{t+1} - g(\Gamma_t)) + \kappa_M g(\Gamma_t) + \lambda c_X \delta(\Gamma_t) g'(\Gamma_t)], \quad (59)$$

over absolutely continuous schedules in $\tilde{\mathcal{G}}$ (defined below).

For each γ , define the conditional coverage-loss function

$$\phi_\gamma(m) := \mathbb{E}[\ell(D_{t+1} - m) | \Gamma_t = \gamma]. \quad (60)$$

If the stationary law of Γ_t admits density f on $(0, \infty)$, then (59) is equivalent to

$$\min_{g \in \tilde{\mathcal{G}}} \int_0^\infty f(\gamma) (\phi_\gamma(g(\gamma)) + \kappa_M g(\gamma)) d\gamma + \lambda c_X \int_0^\infty f(\gamma) \delta(\gamma) g'(\gamma) d\gamma. \quad (61)$$

C.2 Problem solution and the role of slope caps as constraints on g'

Problem (61) separates three forces: (i) higher $g(\gamma)$ reduces expected shortfall at state γ ; (ii) steeper local pass-through $g'(\gamma)$ raises expected update-induced forced-flow costs, weighted by input volatility $\delta(\gamma)$; and (iii) higher levels of margin carry a direct collateral cost. The implementability restriction from the theorem is imposed as a pointwise bound on local pass-through, i.e., a slope cap on g' .

For notational convenience, define

$$w(\gamma) = f(\gamma) \delta(\gamma), \quad (62)$$

so (61) becomes

$$\min_{g \in \tilde{\mathcal{G}}} \int_0^\infty f(\gamma) (\phi_\gamma(g(\gamma)) + \kappa_M g(\gamma)) d\gamma + \lambda c_X \int_0^\infty w(\gamma) g'(\gamma) d\gamma. \quad (63)$$

First-best benchmark

Ignore slope restrictions and take g absolutely continuous with $g(\gamma) \geq 0$. Under mild boundary conditions,

$$\int_0^\infty w(\gamma) g'(\gamma) d\gamma = - \int_0^\infty w'(\gamma) g(\gamma) d\gamma,$$

so the problem is pointwise in γ . Define the effective marginal cost

$$\kappa_M^{\text{eff}}(\gamma) := \kappa_M - \lambda c_X \frac{w'(\gamma)}{f(\gamma)} = \kappa_M - \lambda c_X \frac{(f(\gamma)\delta(\gamma))'}{f(\gamma)}. \quad (64)$$

Any interior first-best schedule g^{fb} satisfies

$$g^{\text{fb}}(\gamma) \in \arg \min_{m \geq 0} \left\{ \phi_\gamma(m) + \kappa_M^{\text{eff}}(\gamma) m \right\}. \quad (65)$$

For $\ell(x) = (x)_+$, $\phi'_\gamma(m) = -\Pr(D_{t+1} > m \mid \Gamma_t = \gamma)$, so when $\kappa_M^{\text{eff}}(\gamma) \in (0, 1)$,

$$\Pr(D_{t+1} > g^{\text{fb}}(\gamma) \mid \Gamma_t = \gamma) = \kappa_M^{\text{eff}}(\gamma), \quad (66)$$

equivalently

$$g^{\text{fb}}(\gamma) = Q_{1-\kappa_M^{\text{eff}}(\gamma)}(D_{t+1} \mid \Gamma_t = \gamma). \quad (67)$$

Under Assumption 16,

$$g^{\text{fb}}(\gamma) = \sqrt{\gamma} Q_{1-\kappa_M^{\text{eff}}(\gamma)}(Y_{t+1}), \quad (68)$$

and when $\lambda = 0$ we recover $\kappa_M^{\text{eff}}(\gamma) \equiv \kappa_M$.

Implementability as a slope cap

Impose implementability as a pointwise bound on pass-through:

$$0 \leq g'(\gamma) \leq s(\gamma) \quad \text{for a.e. } \gamma > 0, \quad (69)$$

with $s(\gamma) = \bar{s}$ (input-scale cap) or $s(\gamma) = \bar{s}/(2\sqrt{\gamma})$ (volatility-scale cap). The feasible set is

$$\tilde{\mathcal{G}} = \{g : \mathbb{R}_+ \rightarrow \mathbb{R}_+ \mid g \text{ absolutely continuous, } g(\gamma) \geq 0, 0 \leq g'(\gamma) \leq s(\gamma) \text{ a.e.}\}. \quad (70)$$

Proposition 4 (Pooling, tracking, and slope saturation). Assume $\phi_\gamma(\cdot)$ is differentiable for a.e. γ . The optimal schedule g^* has three regions: (i) *pooling*, where $g^{*'}(\gamma) = 0$; (ii) *tracking*, where $0 < g^{*'}(\gamma) < s(\gamma)$ and g^* locally satisfies the first-best condition; and (iii) *slope saturation*, where $g^{*'}(\gamma) = s(\gamma)$.

Closed form

$s/(2\sqrt{\gamma})$, so on any slope-saturated region,

$$g^*(\gamma) = A + \bar{s} \sqrt{\gamma} \quad \text{on that region,} \quad (71)$$

with the constant A pinned down by continuity at the boundary where the slope-saturated segment begins. Because the update-cost term in (61) is linear in g' , removing the upper bound in (69) creates an incentive to concentrate adjustment into arbitrarily narrow regions of γ (approximating jumps). Imposing the implementability cap rules out such vanishing-width adjustment and yields the pooling/tracking/saturation structure in Proposition 4. This provides a continuous-state counterpart

to common rulebook tools: pooling corresponds to buffers/floors that locally mute pass-through, tracking corresponds to target-like behavior in intermediate states, and slope saturation corresponds to limited pass-through in stress states where the incentive restriction binds.

C.3 Portfolio margining

We extend the margin-design problem to portfolio margining with cross-impact. There are $K \geq 2$ cleared contracts. At CCP update times, transaction prices satisfy

$$P_t = V_t + A Q_t, \quad Q_t = \Delta X_t + a_t, \quad (72)$$

where $A \in \mathbb{R}^{K \times K}$ is symmetric positive definite, $a_t \in \mathbb{R}^K$ is the strategic (trader) order flow, and $\Delta X_t \in \mathbb{R}^K$ is constrained-sector net order flow. The CCP computes a *scalar* portfolio risk statistic and posts a *scalar* initial margin requirement:

$$\Gamma_t = \Gamma(P_{t-m:t}), \quad M_{t+1} = g(\Gamma_t). \quad (73)$$

When a margin increase binds, constrained investors are forced to deleverage. We summarize the incremental liquidation induced by a scalar margin change by a (state-dependent) liquidation direction $B_t \in \mathbb{R}^K$:

$$\Delta X_{t+1} = \Delta X_{t+1}^{(0)} - B_t \Delta M_{t+1}, \quad \Delta M_{t+1} \equiv M_{t+1} - M_{t+1}^{(0)}. \quad (74)$$

Equivalently, the margin-update-induced flow at $t + 1$ is

$$v_{t+1} \equiv \Delta X_{t+1} - \Delta X_{t+1}^{(0)} = -B_t (M_{t+1} - M_t) = -B_t (g(\Gamma_t) - g(\Gamma_{t-1})). \quad (75)$$

The liquidation-driven component of the next-period mark change is therefore

$$\Delta P_{t+1}^{liq} = A v_{t+1} = -A B_t (g(\Gamma_t) - g(\Gamma_{t-1})). \quad (76)$$

The vector $A B_t$ is the price-pressure direction generated by meeting a higher scalar call. It combines (i) which contracts are liquidated when constraints bind (encoded by B_t) and (ii) how liquidation spills across contracts through cross-impact (encoded by off-diagonal entries of A).

Let $D_{t+1} \geq 0$ denote next-period losses the CCP aims to cover. For a rule g , define the per-period cost

$$L_t(g) := \ell(D_{t+1} - g(\Gamma_t)) + \lambda \mathcal{C}(-B_t (g(\Gamma_t) - g(\Gamma_{t-1}))) + \kappa_M g(\Gamma_t), \quad (77)$$

where $\ell : \mathbb{R} \rightarrow \mathbb{R}_+$ is convex and nondecreasing (coverage loss), $\mathcal{C} : \mathbb{R}^K \rightarrow \mathbb{R}_+$ is convex with $\mathcal{C}(0) = 0$ (market-functioning/procyclicality cost), $\lambda \geq 0$ scales procyclicality costs, and $\kappa_M > 0$ is a level cost of posted margin.

To obtain the same smooth functional structure used in the single-asset design problem, we adopt

a convex, positively homogeneous proxy for the cost of liquidation-induced price pressure:

$$\mathcal{C}(v) = \|Av\|_2. \quad (78)$$

Under (78) and (75),

$$\mathcal{C}(-B_t(g(\Gamma_t) - g(\Gamma_{t-1}))) = \|AB_t\|_2 |g(\Gamma_t) - g(\Gamma_{t-1})|.$$

Assume g is absolutely continuous, with $g'(\gamma) \geq 0$ a.e. (monotonicity). For one-step changes in the input that are small relative to the relevant scale, use the same local linearization as in the single-asset case:

$$g(\Gamma_t) - g(\Gamma_{t-1}) \approx g'(\Gamma_t) (\Gamma_t - \Gamma_{t-1}). \quad (79)$$

Combining $g'(\Gamma_t) \geq 0$ with (79) yields the approximation

$$|g(\Gamma_t) - g(\Gamma_{t-1})| \approx g'(\Gamma_t) |\Gamma_t - \Gamma_{t-1}|.$$

Define the *impact-weighted* one-step input volatility under portfolio liquidation:

$$\delta^{PM}(\gamma) := \mathbb{E}[\|AB_t\|_2 |\Gamma_t - \Gamma_{t-1}| \mid \Gamma_t = \gamma]. \quad (80)$$

Then the procyclicality term admits the smooth approximation

$$\mathbb{E}[\mathcal{C}(-B_t(g(\Gamma_t) - g(\Gamma_{t-1})))] \approx \mathbb{E}[\delta^{PM}(\Gamma_t) g'(\Gamma_t)]. \quad (81)$$

As in the single-asset design section, define for each γ the conditional expected coverage loss

$$\phi_\gamma(m) := \mathbb{E}[\ell(D_{t+1} - m) \mid \Gamma_t = \gamma]. \quad (82)$$

Assume Γ_t admits a stationary density f on $(0, \infty)$, so $dF(\gamma) = f(\gamma) d\gamma$. Under (81), the stationary one-period objective becomes the continuous functional program

$$\min_{g \in \tilde{\mathcal{G}}} \int_0^\infty f(\gamma) (\phi_\gamma(g(\gamma)) + \kappa_M g(\gamma)) d\gamma + \lambda \int_0^\infty f(\gamma) \delta^{PM}(\gamma) g'(\gamma) d\gamma, \quad (83)$$

where $\tilde{\mathcal{G}}$ collects nonnegative, absolutely continuous, weakly increasing schedules satisfying the implementability slope restriction stated below.

Define the portfolio-design weight

$$w^{PM}(\gamma) = f(\gamma) \delta^{PM}(\gamma). \quad (84)$$

Then (83) can be written compactly as

$$\min_{g \in \tilde{\mathcal{G}}} \int_0^\infty f(\gamma) (\phi_\gamma(g(\gamma)) + \kappa_M g(\gamma)) d\gamma + \lambda \int_0^\infty w^{PM}(\gamma) g'(\gamma) d\gamma. \quad (85)$$

Temporarily ignore slope restrictions and assume boundary conditions such that the integration-by-parts identity

$$\int_0^\infty w^{PM}(\gamma) g'(\gamma) d\gamma = - \int_0^\infty (w^{PM})'(\gamma) g(\gamma) d\gamma$$

holds. Then the objective (85) can be rewritten as a pure level integral:

$$\kappa_M^{eff, PM}(\gamma) = \kappa_M - \lambda \frac{(w^{PM})'(\gamma)}{f(\gamma)} = \kappa_M - \lambda \frac{(f(\gamma)\delta^{PM}(\gamma))'}{f(\gamma)}. \quad (86)$$

Any interior first-best schedule g^{fb} therefore solves the pointwise problem

$$g^{fb}(\gamma) \in \arg \min_{m \geq 0} \left\{ \phi_\gamma(m) + \kappa_M^{eff, PM}(\gamma) m \right\}. \quad (87)$$

For the benchmark shortfall loss $\ell(x) = (x)_+$, we have $\partial_m \phi_\gamma(m) = -\mathbb{P}(D_{t+1} > m \mid \Gamma_t = \gamma)$ whenever ϕ_γ is differentiable, so on interior states with $\kappa_M^{eff, PM}(\gamma) \in (0, 1)$,

$$\mathbb{P}(D_{t+1} > g^{fb}(\gamma) \mid \Gamma_t = \gamma) = \kappa_M^{eff, PM}(\gamma), \quad g^{fb}(\gamma) = Q_{1-\kappa_M^{eff, PM}(\gamma)}(D_{t+1} \mid \Gamma_t = \gamma). \quad (88)$$

Portfolio margining admits the same trigger-and-reverse incentive restriction as in the single-asset case, but the relevant local amplification depends on how the risk-increasing direction in mark space aligns with the liquidation-driven price-pressure direction AB_t . Fix the history $P_{t-m:t-1}$ and consider a perturbation of the last mark P_t . Assume there exists a unit direction $d_t \in \mathbb{R}^K$ and a constant $c_{\Gamma, t} > 0$ such that for small perturbations ΔP_t with $d_t^\top \Delta P_t \geq 0$,

$$\Delta \Gamma_t \geq c_{\Gamma, t} d_t^\top \Delta P_t.$$

Define the alignment coefficient

$$\psi_t \equiv \max\{0, -d_t^\top AB_t\}. \quad (89)$$

A sufficient local restriction for ruling out the liquidation-direction trigger-and-reverse deviation is

$$g'(\Gamma_t^{(0)}) \leq s_t^{multi} \equiv \frac{2}{c_{\Gamma, t} \psi_t}, \quad (\psi_t > 0), \quad (\text{ICm})$$

with the convention that the restriction is slack when $\psi_t = 0$. For implementation, we impose a conservative *global* Lipschitz cap stated in terms of the scalar input alone:

$$0 \leq g(\gamma_2) - g(\gamma_1) \leq \bar{s}^{multi}(\gamma_2 - \gamma_1), \quad 0 \leq \gamma_1 \leq \gamma_2, \quad (90)$$

for some constant \bar{s}^{multi} satisfying $\bar{s}^{multi} \leq \inf_t s_t^{multi}$ over the states the CCP aims to protect.

Under monotonicity and the slope cap (90), the constrained optimum g^* retains the same qualitative structure as in the single-asset problem:

- *Pooling (local insensitivity)*: regions where $g^*(\gamma) = 0$.
- *Tracking*: regions where $0 < g^*(\gamma) < \bar{s}^{multi}$ and g^* satisfies the pointwise FOC associated with (87) (equivalently, the quantile rule (88) for $\ell(x) = (x)_+$).

- *Slope saturation*: regions where $g^*(\gamma) = \bar{s}^{multi}$.

Portfolio margining shifts these regions relative to contract-by-contract margining through two channels: (i) the procyclicality weight $\delta^{PM}(\gamma)$ in (80) scales with $\|AB_t\|_2$, so cross-impact and concentrated liquidation increase the marginal cost of short-horizon pass-through in (83); (ii) the admissible slope \bar{s}^{multi} is tightened when ψ_t is large, which can occur precisely when the trigger direction (what moves Γ_t) differs from the liquidation direction (what is sold when the scalar call binds).

Under asset-by-asset margining, the CCP posts a vector of requirements

$$M_{k,t+1} = g_k(\Gamma_{k,t}), \quad k = 1, \dots, K,$$

each computed from a contract-specific risk input $\Gamma_{k,t}$. The key simplification is that, when a contract- k call binds, incremental forced liquidation is primarily in contract k , so the liquidation direction $B_t^{(k)}$ is typically aligned with the trigger direction for $\Gamma_{k,t}$. In that case, the alignment coefficient governing the sufficient IC slope cap depends mainly on the diagonal impact of contract k , whereas under portfolio margining the coefficient ψ_t in (89) can load directly on cross-impact terms because the trigger leg and liquidation leg need not coincide.

C.4 Practical differences between asset-by-asset and portfolio margining

Example 1 (procyclicality and spillovers under cross-impact). Let $K = 2$ and cross-impact be

$$A = \begin{pmatrix} \alpha_1 & \beta \\ \beta & \alpha_2 \end{pmatrix}, \quad \alpha_1 > 0, \alpha_2 > 0, \beta^2 < \alpha_1 \alpha_2.$$

Suppose that when the scalar portfolio call binds, constrained investors deleverage primarily by selling contract 2:

$$B_t = \begin{pmatrix} 0 \\ b \end{pmatrix}, \quad b > 0, \quad \Rightarrow \quad AB_t = b \begin{pmatrix} \beta \\ \alpha_2 \end{pmatrix}.$$

Under the smooth objective (83), the procyclicality term is proportional to $\|AB_t\|_2 |g(\Gamma_t) - g(\Gamma_{t-1})|$, and the corresponding weight satisfies $\delta^{PM}(\gamma) = \mathbb{E}[\|AB_t\|_2 |\Gamma_t - \Gamma_{t-1}| \mid \Gamma_t = \gamma]$. Holding the one-step input move fixed, larger cross-impact spillovers (larger $|\beta|$) raise $\|AB_t\|_2 = b\sqrt{\beta^2 + \alpha_2^2}$: even if liquidation is concentrated in contract 2, the induced price pressure spills into contract 1 through β , increasing the marginal cost of pass-through in volatile states.

Set $\alpha_1 = \alpha_2 = 1$ and $\beta = 0.3$. Then $\|AB_t\|_2 \approx 1.044b$. If the call is met by selling $b = 10$ units per unit margin, the liquidation-induced price-pressure magnitude per unit margin is about 10.44. Under asset-by-asset margining, the liquidation direction (and thus which contract bears the direct forced flow) changes; magnitudes differ only insofar as the liquidation intensity b differs across the two implementations.

Example 2 (incentive slope caps tighten under portfolio wedge). Keep the same A and B_t , and suppose that locally pushing $P_{1,t}$ downward increases the scalar input, so take the risk-increasing

direction $d_t = -e_1$. Then

$$\psi_t = \max\{0, -d_t^\top AB_t\} = \max\{0, b\beta\} = b\beta.$$

If $c_{\Gamma,t} = 1$, the sufficient IC slope bound (ICm) is

$$g'(\Gamma_t^{(0)}) \leq \frac{2}{b\beta}.$$

With $\beta = 0.3$ and $b = 10$, this gives $g'(\Gamma_t^{(0)}) \leq 2/3 \approx 0.667$. Contrast asset-by-asset margining for contract 1. If a binding contract-1 call induces liquidation mainly in contract 1, take $B_t^{(1)} = (b_1, 0)^\top$ and $d_t^{(1)} = -e_1$. Then $\psi_t^{(1)} = \max\{0, -d_t^{(1)\top} AB_t^{(1)}\} = b_1\alpha_1$. With $\alpha_1 = 1$ and $b_1 = 1$, the sufficient bound becomes $g'_1(\Gamma_{1,t}^{(0)}) \leq 2$, which is three times looser than the portfolio bound in the numerical example. The difference is structural: portfolio margining allows the trigger leg (contract 1 moves the scalar input) and the liquidation leg (contract 2 is sold to meet the scalar call) to differ, and cross-impact transmits liquidation pressure back into the trigger contract.

Portfolio margining typically lowers *levels* through netting but can raise the marginal cost of short-horizon adjustments and tighten implementable pass-through (Examples 1 and 2). In the design problem, this shifts the optimal schedule toward more pooling and flatter adjustment regions, unless the CCP compensates by raising a floor or buffer (increasing levels in calm states) to preserve coverage.

C.5 Why jumps are not a fix under portfolio margining

A natural instinct is to replace steep slopes by discrete steps. In this environment, upward jumps do not solve the incentive problem; they replace smooth marginal manipulation by threshold manipulation. If the intended input is arbitrarily close below a jump threshold, an arbitrarily small trade can cross the threshold, trigger a discrete increase in the next call, and thereby induce a discrete liquidation event. When constraints bind, this creates a predictable price component that a short-horizon round trip can monetize.

Lemma 17 (Jump discontinuities create a threshold-crossing manipulation under portfolio margining). *Suppose g is weakly increasing and has an upward jump of size $\Delta M > 0$ at some γ^\dagger , meaning that $\lim_{\gamma \uparrow \gamma^\dagger} g(\gamma) + \Delta M = g(\gamma^\dagger)$. Fix a state in which (i) the risk input is locally manipulable so that an admissible trigger trade can raise Γ_t from $\gamma^\dagger - \varepsilon$ to at least γ^\dagger for arbitrarily small $\varepsilon > 0$; (ii) when the scalar call binds, the jump ΔM induces incremental forced liquidation $v = -B_t\Delta M$; and (iii) prices have cross-impact matrix A . If there exists a unit direction $u \in \mathbb{R}^K$ (a trade direction) such that $u^\top AB_t > 0$ and*

$$u^\top AB_t \Delta M > 2\tau, \tag{91}$$

then for states with $\Gamma_t^{(0)}$ arbitrarily close below γ^\dagger there exists an admissible two-period round trip that crosses the threshold and yields strictly positive conditional expected profit net of proportional execution cost τ .

Proof. Fix $\varepsilon > 0$ and consider a state with $\Gamma_t^{(0)} = \gamma^\dagger - \varepsilon$. By assumption, the trader can choose a trigger trade at time t that raises the realized input to at least γ^\dagger , triggering the jump. The jump implies a

margin increase of at least ΔM and therefore incremental forced liquidation $v = -B_t \Delta M$ at $t + 1$, which shifts the transaction price by the liquidation-driven component $\Delta P_{t+1}^{liq} = -AB_t \Delta M$.

Consider the two-period round trip that sells qu at time t and buys qu at time $t + 1$, for some $q > 0$ small enough to remain admissible. The round trip profit equals $qu^\top (P_t - P_{t+1})$ minus proportional execution costs $2\tau q$. Under the maintained no-drift condition for fundamentals and baseline flows, the predictable component of $u^\top (P_t - P_{t+1})$ comes from liquidation, contributing approximately $u^\top AB_t \Delta M$. The trader's own round-trip impact loss is $O(q)$ in this normalization, so for q small enough the sign of expected profit is governed by $u^\top AB_t \Delta M - 2\tau$. Under (91), this is strictly positive, so the deviation yields strictly positive conditional expected profit. Since ε can be arbitrarily small, the trigger needed to cross the threshold can be made arbitrarily small, so the deviation is admissible in states arbitrarily close below γ^\dagger . \square

Lemma 17 shows why steps are not an improvement relative to pooling-plus-adjustment. A jump makes the mapping from the risk input to posted margin discontinuous, so an arbitrarily small movement in Γ_t can create a discrete increase in M_{t+1} . Replacing the jump by an adjustment region of finite width in Γ , over which margin increases at a bounded slope, removes this vanishing-cost deviation. This is exactly what pooling-plus-adjustment implements.

D Calibration details for Section 3.6

This appendix documents the empirical inputs used to interpret the loop-gain magnitude objects in Section 3.6.

D.1 Objects and identities

Fix the one-update trigger-and-reverse deviation from Lemma 5: sell $q > 0$ at the sampling time t and buy q at $t + 1$. Under linear impact, selling q distorts the sampled mark by $\Delta P_t = \alpha q$. Define

$$\delta_{\text{bps}} = 10^4 \frac{\alpha q}{P_t}, \quad G = \frac{L}{q}, \quad m = \frac{M_t}{P_t}.$$

Ignoring proportional costs and inventory penalties ($\kappa = \tau = \rho = 0$), Lemma 5 implies the bps profit and ROI lower bounds

$$\pi_{\text{bps}} = 10^4 \frac{\mathbb{E}[\Pi_1 | S_t]}{q P_t} \geq (G - 2) \delta_{\text{bps}}, \quad ROI\% = 100 \frac{\mathbb{E}[\Pi_1 | S_t]}{m q P_t} \geq \frac{(G - 2) \delta_{\text{bps}}}{100 m}. \quad (92)$$

The liquidation-driven price pressure satisfies $|\Delta P_{t+1}^{\text{liq}}| = \alpha L = G \alpha q$, so in bps the induced liquidation pressure is $G \delta_{\text{bps}}$.

D.2 Impact

Let ADV denote average daily traded value (dollars) and let $N_q = q P_t$ denote the trigger notional. Define the participation rate

$$f = \frac{N_q}{ADV} \in (0, 1].$$

Empirical microstructure evidence often supports square-root scaling of costs in participation units. A convenient benchmark is

$$\delta_{\text{bps}}(f) \approx 200 \sqrt{f}, \quad (93)$$

where f is expressed as a fraction of full-day ADV . The constant 200 corresponds to a one-way cost scale of roughly $100\sqrt{f}$ bps combined with the standard triangle heuristic that maps average execution cost into a peak mark distortion at the sampling time. Using full-day ADV is conservative for settlement and margining since the relevant sampling-window volume is typically smaller than full-day volume, which raises effective participation and thus δ_{bps} . For instance, from equation (93) we have:

$$f = 1\% \Rightarrow \delta_{\text{bps}} \approx 20, \quad f = 2\% \Rightarrow \delta_{\text{bps}} \approx 28, \quad f = 5\% \Rightarrow \delta_{\text{bps}} \approx 45, \quad f = 10\% \Rightarrow \delta_{\text{bps}} \approx 63.$$

Inverting (93) gives $f \approx (\delta_{\text{bps}}/200)^2$, so even very small bps distortions correspond to extremely small fractions of full-day ADV when the sampling window is thin. This is consistent with Frazzini et al. (2018) and Kyle and Obizhaeva (2016).

D.3 Elasticities

Let X denote the position of the constraint-sensitive sector. Define the elasticity of this sector's position to the margin level M by

$$\varepsilon_{X,M} = -\frac{d \log X}{d \log M}.$$

For a relative increase $\Delta M/M$, the implied fractional position reduction is

$$\frac{L}{X} \approx \varepsilon_{X,M} \frac{\Delta M}{M}. \quad (94)$$

Empirical work on futures margins reports economically meaningful elasticities. To keep the calibration conservative and avoid selecting extreme estimates, we use the range

$$\varepsilon_{X,M} \in \{0.15, 0.225, 0.30\}.$$

This range is consistent with evidence that margin increases reduce open interest and speculative demand over horizons relevant for margin adjustment, e.g., the evidence in [Hedegaard \(2014\)](#).

D.4 Forced-liquidation multiple

Define the turnover ratio

$$\tau = \frac{X P_t}{ADV}.$$

Since $f = (q P_t)/ADV$, we have $q/X = f/\tau$. Combining with (94) yields

$$G = \frac{L}{q} \approx \varepsilon_{X,M} \frac{\Delta M}{M} \frac{\tau}{f}. \quad (95)$$

This identity shows that G increases when: (i) margin moves are large ($\Delta M/M$ high), (ii) the constrained sector is large relative to market depth (τ high), and (iii) the trigger is a small share of volume (f low).

D.5 Margin increases

Section 3.6 uses relative margin jumps

$$\frac{\Delta M}{M} \in \{20\%, 40\%, 60\%\},$$

which correspond to moving from a baseline margin ratio $m_0 = 5\%$ to $m_1 \in \{6\%, 7\%, 8\%$ under the local linearization $m := M/P$. Risk-based initial margin models can generate short-horizon margin calls of this order even with anti-procyclicality tools. Stress-episode evidence and model-based simulations document large and discrete changes in initial margin over short horizons; these facts motivate using 20% to 60% as routine stressed and binding scenarios rather than tail hypotheticals.

D.6 Tables

Table 2 Calibration inputs used in Appendix D.

Object	Symbol	Values
Baseline margin ratio	m_0	5%
Post-jump margin ratio	m_1	6%, 7%, 8%
Relative margin jump	$\Delta M/M$	20%, 40%, 60%
Position-margin elasticity	$\varepsilon_{X,M}$	0.15, 0.225, 0.30
Turnover ratio	τ	0.5, 1, 2
Participation rate	f	0.5%, 1%, 2%, 5%
Mark distortion mapping	$\delta_{\text{bps}}(f)$	$200\sqrt{f}$

Table 3

Forced-liquidation multiple $G = L/q$ implied by (95). Each cell lists G for $\varepsilon_{X,M} = 0.15/0.225/0.30$.

f	$\tau = XP_t/ADV$		
	0.5	1	2
Panel A: $\Delta M/M = 20\%$			
0.5%	3.0/4.5/6.0	6.0/9.0/12.0	12.0/18.0/24.0
1%	1.5/2.25/3.0	3.0/4.5/6.0	6.0/9.0/12.0
2%	0.75/1.125/1.5	1.5/2.25/3.0	3.0/4.5/6.0
5%	0.30/0.45/0.60	0.60/0.90/1.20	1.20/1.80/2.40
Panel B: $\Delta M/M = 40\%$			
0.5%	6.0/9.0/12.0	12.0/18.0/24.0	24.0/36.0/48.0
1%	3.0/4.5/6.0	6.0/9.0/12.0	12.0/18.0/24.0
2%	1.5/2.25/3.0	3.0/4.5/6.0	6.0/9.0/12.0
5%	0.60/0.90/1.20	1.20/1.80/2.40	2.40/3.60/4.80
Panel C: $\Delta M/M = 60\%$			
0.5%	9.0/13.5/18.0	18.0/27.0/36.0	36.0/54.0/72.0
1%	4.5/6.75/9.0	9.0/13.5/18.0	18.0/27.0/36.0
2%	2.25/3.375/4.5	4.5/6.75/9.0	9.0/13.5/18.0
5%	0.90/1.35/1.80	1.80/2.70/3.60	3.60/5.40/7.20

Table 4

ROI lower bound in percent using (92) with $\delta_{\text{bps}}(f)$ from (93) and the conservative denominator $m_1 q P_t$. Each cell lists ROI% for $\varepsilon_{X,M} = 0.15/0.225/0.30$. Negative entries correspond to $G < 2$.

f	$\tau = X P_t / \text{ADV}$		
	0.5	1	2
Panel A: $m_0 = 5\% \rightarrow m_1 = 6\%$ ($\Delta M/M = 20\%$)			
0.5%	2.8/7.1/11.4	11.4/19.9/28.5	28.5/45.6/62.6
1%	-2.0/1.0/4.0	4.0/10.1/16.1	16.1/28.2/40.3
2%	-7.1/-5.0/-2.8	-2.8/1.4/5.7	5.7/14.2/22.8
5%	-15.3/-14.0/-12.6	-12.6/-9.9/-7.2	-7.2/-1.8/3.6
Panel B: $m_0 = 5\% \rightarrow m_1 = 7\%$ ($\Delta M/M = 40\%$)			
0.5%	9.8/17.1/24.4	24.4/39.0/53.7	53.7/83.0/112.3
1%	3.5/8.6/13.8	13.8/24.2/34.5	34.5/55.2/75.9
2%	-2.4/1.2/4.9	4.9/12.2/19.5	19.5/34.2/48.8
5%	-10.8/-8.5/-6.2	-6.2/-1.5/3.1	3.1/12.3/21.6
Panel C: $m_0 = 5\% \rightarrow m_1 = 8\%$ ($\Delta M/M = 60\%$)			
0.5%	14.9/24.6/34.2	34.2/53.4/72.6	72.6/111.0/149.5
1%	7.5/14.3/21.1	21.1/34.7/48.3	48.3/75.5/102.7
2%	1.1/5.9/10.7	10.7/20.3/29.9	29.9/49.1/68.3
5%	-7.4/-4.4/-1.4	-1.4/4.7/10.8	10.8/23.0/35.1

E How CCP margin engines map into the model

Real-world initial margin systems differ in engineering detail, but their architecture is stable and is described in rulebooks, methodology notes, and parameter files. In most implementations, a core engine produces a scalar risk number from current marks and a prescribed set of risk scenarios or recent price history. We denote this core input by Γ_t . An overlay layer then applies anti-procyclicality tools, floors, add-ons, rounding, and governance adjustments, producing the posted requirement

$$M_{t+1} = g(\Gamma_t).$$

Our model is not a claim that any one CCP uses realized variance. It is a claim about the composite mapping from sampled, transaction-based marks into posted margin, and about what happens when that mapping is locally sensitive in states where initial margin constraints bind.

Two clarifications matter once we allow portfolio margining, as in Section 5. First, the marks are naturally a vector, even for a single clearing service: different contracts, maturities, and options share a portfolio margin system. The core statistic Γ_t is then a portfolio object computed from a vector of marks and from portfolio positions, while M_{t+1} remains a single scalar requirement. Second, a binding call is a funding event. Meeting it can force liquidation in a subset of contracts that need not coincide with the subset of contracts that most efficiently moves the statistic. This is why the multi-asset extension introduces a liquidation direction: it summarizes, locally, how a scalar margin change turns into a vector of forced trades. The mechanism in the paper does not require that the CCP explicitly chooses this liquidation direction; it only requires that, when constraints bind, a higher posted requirement generates nontrivial forced deleveraging somewhere in the portfolio.

The impossibility result is local and it separates what matters. If the risk input Γ_t can be moved at first order by trading inside the sampling grid used by the CCP, and if the posted mapping is locally risk-sensitive where constraints bind, then we face a design tension. Either there exist profitable short-horizon round trips that work through the margin update, or the system suppresses local sensitivity by inserting non-smooth features (max operators, thresholds, rounding, discrete add-ons) that concentrate adjustment at kinks and can create liquidity cliffs. The examples below show how standard methodologies map into (Γ_t, g) .

E.1 SPAN-style margining: CME SPAN and other venues

We start from SPAN (Standard Portfolio Analysis of Risk) because its methodology is unusually well documented, both in a formal methodology note and in public reference documentation describing the risk parameter files and related inputs (CME Group, 2019c,b, 2026). In broad terms, SPAN is a market-simulation Value-at-Risk system: it evaluates a portfolio by computing gains and losses under a finite set of standardized risk scenarios over a specified horizon (typically one trading day) and then applying additional charges and credits that capture spread risk, delivery risk, and residual option risk (CME Group, 2019c,b).

Although SPAN is often associated with CME, it is used more broadly across exchanges and clearing organizations (CME Group, 2019c,b). For example, LME Clear uses SPAN for initial margin and

publishes SPAN margin parameter files and a public SPAN methodology document that describes the same scenario-based structure, combined-commodity grouping, and the subsequent spread and option-minimum layers (London Metal Exchange, 2026a, 2015, 2026b). This cross-venue use matters for our purposes: the mapping from marks into a scenario-based risk number and then into posted margin is not a CME-specific implementation detail.

To map SPAN into our notation, fix an update time t and collect the public SPAN parameters (risk scenarios, scan ranges, spread tables, delivery parameters, short-option-minimum parameters, and related conventions) into Θ_t . SPAN first groups instruments by underlying into combined commodities (CME Group, 2019c,b). Within a combined commodity, each contract is associated with a risk array, a list of scenario profit-and-loss values obtained by moving the underlying price and (for options) implied volatility over ranges set by the margin authority (CME Group, 2019c,b). Most SPAN implementations use 16 scenarios, but the number is a design choice (CME Group, 2019c). Let S denote the scenario set, and let $l_s(P_t; \Theta_t)$ be the portfolio loss under scenario s computed from the prevailing marks and the risk arrays. The scan-risk object is the worst loss across scenarios,

$$\Gamma_t^{\text{scan}} := \max_{s \in S} l_s(P_t; \Theta_t).$$

This is already a pointwise maximum and therefore piecewise smooth in marks, with kinks when the identity of the worst scenario changes.

The posted requirement is not only Γ_t^{scan} . SPAN adds layers that are mechanical and often non-smooth. In the CME documentation, the combined-commodity requirement compares (scan risk plus an intra-commodity spread charge and delivery risk, net of an inter-commodity spread credit) to a short option minimum, and takes the larger number (CME Group (2019c,b)). In addition, the total portfolio requirement adjusts for option premium (net option value) and applies rounding and minimum conventions (CME Group (2019c)). LME Clear presents the same structure in SPAN terms: scanning risk, inter-prompt spread charges, inter-contract credits, and a short option minimum floor, with parameter files generated on a daily schedule (including intraday and end-of-day generation) (London Metal Exchange (2015, 2026a)). We can represent these layers by writing

$$M_{t+1} = g(\Gamma_t^{\text{scan}}; \Theta_t),$$

where g collects spread charges and credits, delivery add-ons, the short option minimum, option-value adjustments, rounding, and any other overlays.

This structure interacts naturally with the multi-asset mechanism. SPAN is portfolio margining: Γ_t^{scan} is computed from a vector of marks and from portfolio aggregation, while M_{t+1} is a single scalar requirement. When the requirement binds, the forced deleveraging needed to meet the higher requirement is a portfolio response. It can be concentrated in specific legs, for example because some positions are more liquid, more easily reduced, or more binding under internal constraints. In the notation of Section 5, this is exactly the role of the liquidation direction: a scalar change in M_{t+1} maps into a vector of forced trades. The cross-contract wedge then becomes relevant: a deviation at time t can be chosen to move the portfolio statistic through the contracts that have the largest effect on Γ_t , while the induced liquidation can hit other contracts where the trader has positioned to profit.

SPAN also generates non-smoothness for two distinct reasons, each of which matters for the impossibility logic. First, Γ_t^{scan} is a maximum over finitely many scenario losses, so it has scenario-switching kinks in marks. Second, the overlay g contains max operators (for the short option minimum), thresholded charges and credits (through spread tables and delivery parameters), and explicit discrete conventions (rounding, minimum charges, and add-on triggers). When these objects load on sampled transaction-based marks at predictable times, a feasible perturbation of a sampled mark can move the active scenario loss, and therefore the composite posted requirement, at first order. If the system hardens against this by relying more heavily on thresholds and discretization, local sensitivity is reduced, but adjustment is concentrated at kinks and can become cliff-like in states where constraints bind. This is exactly the continuity side of the impossibility result.

Finally, we note that CME has also described a next-generation SPAN 2 framework that maintains the core SPAN calculations while adding more granular and dynamic adjustments and additional risk-factor reporting (CME Group, 2019a). We do not analyze SPAN 2 separately here. For our purposes, the relevant objects remain the same: a scenario-based risk input built from marks and a mapping that translates it into posted margin, potentially with more state dependence and more layers.

E.1.1 OCC STANS

OCC calculates margin at the portfolio level using STANS (System for Theoretical Analysis and Numerical Simulations). Public descriptions emphasize two defining features. First, STANS is a simulation-based methodology: it uses large-scale Monte Carlo simulations of underlying risk factors to generate a distribution of portfolio losses. Second, the base risk measure is a tail functional. OCC describes the base component as 99% Expected Shortfall over a two-day horizon, and it then adds stress and other components as overlays. (The Options Clearing Corporation, 2026b; U.S. Securities and Exchange Commission, 2024b)

In terms of our notation, the natural mapping is directly to the portfolio setting in Section 5. Let q_t denote the portfolio of cleared positions at an update time t and let Z_t denote the collection of marks and risk-factor state variables used to seed the simulation (prices and, for options, implied volatilities are central inputs). Under STANS, simulated risk-factor moves generate simulated portfolio losses over the margin period of risk. Writing these simulated losses as $L^{(n)}(q_t; Z_t)$ for draws $n = 1, \dots, N$, a transparent core statistic is

$$\Gamma_t^{\text{STANS}} := \text{ES}_{0.99}(\{L^{(n)}(q_t; Z_t)\}_{n=1}^N),$$

and the posted requirement is again a composite

$$M_{t+1} = g(\Gamma_t^{\text{STANS}}),$$

where g absorbs the stress test component and other add-ons described in public documentation. OCC's public overview makes the decomposition explicit: a base 99% Expected Shortfall component is adjusted by a stress test component, and the stress test component is itself constructed from multiple sub-components (Dependence and Concentration) with max-type aggregation. (The Options

Clearing Corporation, 2026b; U.S. Securities and Exchange Commission, 2024b)

The risk statistic is computed from a vector of marks and portfolio aggregation, and the output is a scalar requirement. This is exactly the object studied in Section 5: the margin call is scalar, while any forced deleveraging response is a portfolio response. In particular, once the requirement binds, the liquidation needed to meet a higher posted requirement can be concentrated in specific contracts within the portfolio, which is the mechanism summarized by the liquidation direction B_t in (26). We do not need to take a stand on which contracts are liquidated in practice. We only need the local fact that, when the call binds, a higher posted requirement induces nontrivial liquidation somewhere in the portfolio.

From the perspective of the impossibility mechanism, the key object is local sensitivity of the posted requirement to sampled marks. STANS is not written as a function of transaction prices alone, but the simulated loss distribution depends on the marks used to value the portfolio and seed the risk-factor model. When economically dominant components of Z_t are constructed from sampled transaction-based marks at predictable times, a feasible perturbation of a sampled mark can change the simulated loss distribution and therefore change Expected Shortfall at first order, in the sense of Assumption 5 (or Assumption 7 in the multi-asset setting). At the same time, STANS contains several non-smooth ingredients that are natural analogs of the non-smooth overlays we study: for example, OCC’s public description uses max-type conservative choices for volatility inputs and max-type aggregation in stress components, and it applies additional charges in specific situations. These features reduce smooth marginal sensitivity in some regions, but they concentrate adjustment at kinks in the composite mapping from marks to posted margin. (The Options Clearing Corporation, 2026b)

OCC also operates an intraday framework that recalculates STANS-based risk from frequent snapshots and applies escalation criteria when risk increases breach thresholds. SEC-described implementations reference portfolio position sets updated every 20 minutes during regular trading hours (and at least hourly during extended hours), and add-on charges designed to address intraday and overnight risk increases. (U.S. Securities and Exchange Commission, 2024a)

Finally, STANS appears in OCC’s cross-margin programs. In these programs, participating clearinghouses establish joint accounts and compute clearing-level margins from combined positions using STANS. This is a direct real-world instance of the multi-asset perspective in Section 5: a portfolio-level statistic is computed from a vector of marks and positions across products, while a binding call can force liquidation in the subset of contracts that is operationally available to the clearing member. (The Options Clearing Corporation, 2026a)

E.1.2 LCH SwapClear PAIRS

LCH describes PAIRS (Portfolio Approach to Interest Rate Scenarios) as the core initial margin methodology for SwapClear. The public description has two stable elements. First, PAIRS is a filtered-historical-simulation framework with volatility scaling: the model uses a long history of market data to generate a loss distribution by revaluing current positions under scaled historical market moves. Second, the posted risk number is a tail functional. LCH describes PAIRS as an expected shortfall model based on filtered historical simulation incorporating volatility scaling, using ten years of his-

torical market data to simulate changes in portfolio value and estimate a potential loss distribution. (LCH, 2026)

Public filings provide additional operational detail that is useful for mapping PAIRS into our notation. One description states that PAIRS simulates changes to the value of a SwapClear portfolio based on historical market moves from the most recent 2500 business days (approximately ten years), with a confidence level of at least 99.7%. The same description states that the model considers historical foreign exchange rate movements to account for FX risk, and assumes a holding (close-out) period of five days for proprietary positions and seven days for customer positions. The base initial margin requirement is described as the average of the six largest simulated losses generated by the PAIRS model. (LCH.Clearnet Limited, 2015) We use these statements only to pin down the structure of the risk functional: PAIRS is a portfolio revaluation model, the loss distribution is generated from scaled historical moves, and the tail is summarized by an average of extreme losses, which is a discrete expected shortfall-type functional.

The SwapClear procedures also make clear that the posted requirement is composite. PAIRS produces a core statistic, and LCH may apply additional margin components such as a counterparty risk multiplier, liquidity risk margin requirements, tenor basis risk add-ons, and intraday margin calls. (LCH Limited, 2025) The procedures explicitly describe that yield curve scenarios are used to capture potential losses based on observed history, that scenario parameters are monitored and reviewed, and that the Clearing House retains discretion to vary rates for the whole market or for specific accounts. They also describe a liquidity risk margin requirement based on the expected cost of hedging in a default scenario, calibrated using periodic liquidity surveys, and a counterparty risk multiplier margin. (LCH Limited, 2025) Public descriptions outside the rulebook similarly emphasize that, in addition to PAIRS initial margin, LCH applies add-ons covering credit risk and liquidity risk where a particular member’s risk exposure is not captured within the base model. (LCH, 2026) Finally, external assessments report that LCH also uses floor-type tools aimed at procyclicality control, including a floor based on a long unscaled lookback period, and that intraday monitoring can trigger margin calls when liabilities breach predetermined thresholds. (Reserve Bank of Australia, 2019; LCH Ltd, 2022)

This fits the paper’s basic (Γ_t, g) structure. We can write

$$\Gamma_t^{\text{PAIRS}} = (\text{tail functional of simulated/scenario losses}), \quad M_{t+1} = g(\Gamma_t^{\text{PAIRS}}),$$

where g captures the published overlay layer (and, where relevant, the discretionary and governance elements described in the procedures). Because PAIRS is a portfolio model, it maps most directly into the multi-asset setting in Section 5. The inputs to the engine are a vector of marks and risk-factor state variables (yield curves and related variables, potentially across currencies), and the output is a single scalar requirement. When that requirement binds, meeting a higher call is a portfolio response. In our notation, this is exactly the role played by a liquidation direction in (26): a scalar change in M_{t+1} can induce forced trades concentrated in a subset of instruments or curve points. The procedures describe this layer in operational terms (for example via liquidity risk margin requirements and intraday calling practices) rather than as an explicit liquidation rule, and we do not attempt to model that operational choice. We use only the local implication that, when constraints bind, a higher posted

requirement induces nontrivial deleveraging somewhere in the portfolio.

For the impossibility mechanism, the key technical object is local sensitivity of the posted requirement to sampled marks. Here we need to be careful about what is and is not implied by the public descriptions. PAIRS is not presented as a function of transaction prices alone. The risk number depends on the marks and risk-factor series used to value the portfolio and seed the historical revaluation, and those marks may be constructed from curve-building conventions and snapshots that are not fully transparent. For this reason, we do not treat “attackability of PAIRS” as an unconditional claim. What we can establish from the model class is narrower: if some economically dominant input mark used by PAIRS is locally sensitive to tradable prices at predictable sampling times, and if the volatility scaling step loads on recent mark changes with positive weight, then the PAIRS risk number has a first-order channel from the last sampled mark to the posted risk number. Whether a trader can actually exploit this channel is an empirical question about mark construction, market depth, the relevant smoothing parameters, and whether margin constraints bind strongly enough to generate price impact through forced liquidation.

To connect PAIRS more transparently to the single-asset algebra in the main text, we write a stylized one-factor representation that isolates the volatility-scaling channel. We emphasize that this is a reduced-form proxy for one component of a filtered-historical-simulation engine. It is not a claim that the full PAIRS engine is one-dimensional. It captures the local feature we need: the risk number scales with a state-dependent volatility estimate that loads on recent returns. A convenient proxy is an EWMA-type recursion,

$$\hat{\sigma}_t^2 = \lambda \hat{\sigma}_{t-1}^2 + (1 - \lambda) r_t^2, \quad \lambda \in (0, 1), \quad (96)$$

where $r_t = \log P_t - \log P_{t-1}$ is the last sampled log return of the relevant mark P_t . Public descriptions of PAIRS and filings describing PAIRS-style volatility scaling are consistent with this type of exponentially weighted volatility estimate. (LCH, 2026; LCH.Clearnet Limited, 2015)

In a filtered-historical-simulation implementation, the model constructs scaled historical shocks by standardizing past returns and then rescaling them to the current volatility regime. One stylized construction is

$$\tilde{r}_t^{(k)} = \hat{\sigma}_t \varepsilon_{t-k}, \quad \varepsilon_{t-k} = \frac{r_{t-k}}{\hat{\sigma}_{t-k}}, \quad k = 1, \dots, H,$$

where H is the number of historical shocks used.²⁰ For a one-unit long position, define scenario losses by $l_t^{(k)} = -\tilde{r}_t^{(k)}$. The core statistic is then an expected shortfall of the loss sample,

$$\Gamma_t^{\text{PAIRS}} := \text{ES}_p(\{l_t^{(k)}\}_{k=1}^H), \quad M_{t+1} = g(\Gamma_t^{\text{PAIRS}}). \quad (97)$$

Expected shortfall is positively homogeneous, so in this representation Γ_t^{PAIRS} is proportional to $\hat{\sigma}_t$ with proportionality factor $\text{ES}_p(\{-\varepsilon_{t-k}\}_{k=1}^H)$, which is strictly positive as long as the standardized historical sample contains loss outcomes in the tail.

The next proposition records the local implication we use in the general theorem. It is not a claim that a trader can manipulate PAIRS in practice. It is a statement about the mapping in the stylized scaling representation: if the scaling factor loads on the last sampled return with positive weight,

²⁰We use H here to avoid confusion with the number of contracts K in Section 5.

then the risk number inherits a first-order sensitivity to the last sampled mark, holding the rest of the inputs fixed.

Proposition 5 (PAIRS-style volatility scaling implies local manipulability). Fix a reachable state at time t and hold $(P_{t-H}, \dots, P_{t-1})$ fixed. Assume $\lambda \in (0, 1)$ in (96) and suppose that the proportionality factor in (97), namely $\text{ES}_p(\{-\varepsilon_{t-k}\}_{k=1}^H)$, is strictly positive. If $|r_t| \geq R > 0$, then there exist $\delta > 0$ and $c_\Gamma > 0$ such that, for all $0 < |\Delta P| < \delta$ with $\text{sgn}(\Delta P) = \text{sgn}(r_t)$,

$$\Gamma_t^{\text{PAIRS}}(P_t + \Delta P) - \Gamma_t^{\text{PAIRS}}(P_t) \geq c_\Gamma |\Delta P|.$$

Proof. Holding $(P_{t-H}, \dots, P_{t-1})$ fixed, the standardized historical shocks $\{\varepsilon_{t-k}\}_{k=1}^H$ are fixed. In (97), the only dependence of Γ_t^{PAIRS} on P_t is through the current scaling factor $\hat{\sigma}_t$, which depends on $r_t = \log P_t - \log P_{t-1}$ in (96).

From (96),

$$\hat{\sigma}_t^2 = \lambda \hat{\sigma}_{t-1}^2 + (1 - \lambda) r_t^2.$$

Differentiating both sides with respect to P_t (holding P_{t-1} fixed) gives

$$2\hat{\sigma}_t \frac{d\hat{\sigma}_t}{dP_t} = 2(1 - \lambda) r_t \frac{dr_t}{dP_t}.$$

Since $r_t = \log P_t - \log P_{t-1}$, we have $dr_t/dP_t = 1/P_t$. Therefore,

$$\frac{d\hat{\sigma}_t}{dP_t} = \frac{(1 - \lambda) r_t}{\hat{\sigma}_t P_t}.$$

Under $|r_t| \geq R > 0$, this derivative is nonzero at the baseline state. By continuity of the right-hand side in P_t , there exists $\delta > 0$ such that for all $|\Delta P| < \delta$, the derivative keeps the same sign as r_t and satisfies a uniform lower bound in absolute value on the segment between P_t and $P_t + \Delta P$:

$$\left| \frac{d\hat{\sigma}_t}{dP} \Big|_{P=P_t+\theta} \right| \geq c_\sigma \quad \text{for all } \theta \in [-|\Delta P|, |\Delta P|],$$

for some constant $c_\sigma > 0$.

Fix ΔP with $\text{sgn}(\Delta P) = \text{sgn}(r_t)$ and $0 < |\Delta P| < \delta$. By the mean value theorem applied to $P \mapsto \hat{\sigma}_t(P)$, there exists ξ between P_t and $P_t + \Delta P$ such that

$$\hat{\sigma}_t(P_t + \Delta P) - \hat{\sigma}_t(P_t) \geq c_\sigma |\Delta P|.$$

In (97), $l_t^{(k)} = -\hat{\sigma}_t \varepsilon_{t-k}$, so by positive homogeneity of expected shortfall,

$$\Gamma_t^{\text{PAIRS}}(P_t) = \hat{\sigma}_t(P_t) \text{ES}_p(\{-\varepsilon_{t-k}\}_{k=1}^H).$$

By assumption, the factor $\text{ES}_p(\{-\varepsilon_{t-k}\}_{k=1}^H)$ is strictly positive. Hence,

$$\Gamma_t^{\text{PAIRS}}(P_t + \Delta P) - \Gamma_t^{\text{PAIRS}}(P_t) = (\hat{\sigma}_t(P_t + \Delta P) - \hat{\sigma}_t(P_t)) \text{ES}_p(\{-\varepsilon_{t-k}\}_{k=1}^H) \geq c_\Gamma |\Delta P|,$$

where we set $c_\Gamma = c_\sigma \cdot \text{ES}_p(\{-\varepsilon_{t-k}\}_{k=1}^H) > 0$. □

Proposition 5 provides the risk-input sensitivity condition used in our general theorem, but it should be read as a local statement about a volatility-scaling component holding the rest of the inputs fixed. Turning this sensitivity into an economically exploitable round trip requires the additional ingredients emphasized in the paper: a locally risk-sensitive pass-through in the composite mapping g , binding constraints that generate forced liquidation, and price impact. PAIRS is also implemented with overlays and anti-procyclicality tools, including add-ons, intraday calling practices, and floor-type features documented in procedures and external assessments. (LCH Limited, 2025; Reserve Bank of Australia, 2019; LCH Ltd, 2022) These features can attenuate local smooth sensitivity in some regions, but they also concentrate adjustment into non-smooth events, which is the continuity side of the impossibility mechanism.

E.1.3 European anti-procyclicality measures as transformations inside g

In Europe, EMIR technical standards require CCPs to address margin procyclicality. Commission Delegated Regulation (EU) No 153/2013, Article 28, requires a CCP to employ at least one of three measures: an anti-procyclicality buffer, stressed weights, or a long-lookback floor. (European Union, 2013; European Securities and Markets Authority, 2018) Subsequent ESMA work documents heterogeneity in implementation and discusses refinements to these tools. (European Securities and Markets Authority, 2022, 2023) In our notation, these measures enter as deterministic transformations inside the composite mapping g , because they take a core risk number (whatever the CCP uses as Γ_t) and then apply an overlay rule before posting M_{t+1} .

The specific design elements map cleanly into the objects we have already isolated. Floors and buffers introduce regions where marginal pass-through from the risk number to posted margin is reduced or eliminated. When these tools bind and then stop binding, they generate kinks in the mapping from Γ_t to M_{t+1} . If a release rule is itself threshold-based (for example, a buffer that is released only after a discrete trigger), then the composite mapping can have jump-like behavior at the release threshold. Stressed weights work differently: they dampen sensitivity to calm-period data by mixing in stressed-period information, but they can introduce regime-switching behavior when the stressed component becomes active or when the weights change. The theory does not treat any one of these tools as exceptional. They are all part of the same object, g , because the economic mechanism depends on the composite mapping from sampled marks to posted margin, not on how the mapping is labeled in the rulebook.

E.2 Motivating institutional context

Central counterparties (CCPs) are a canonical implementation of price-based rulebooks: post-crisis reforms moved a large share of standardized derivatives into central clearing, where published methodologies map recent marks into variation margin cash flows and initial margin requirements.²¹ A CCP novates trades and manages counterparty default risk with variation margin, initial margin,

²¹See, among the others, Duffie and Zhu (2011), Menkveld and Vuillemeij (2021), Commodity Futures Trading Commission (2026b), European Central Bank (2021).

and a default waterfall. Contracting theories that treat the CCP as an agent clarify how default-waterfall layers and CCP capital can be pinned down by monitoring incentives and member bargaining; see [Kuong and Maurin \(2024\)](#). Central clearing reshapes exposures through netting, and collateral must be posted in cash or high-quality liquid assets, so margins can transform illiquid positions into liquid collateral and vice versa and thereby operate as binding funding constraints in stress states.²² When margin calls are large and correlated, they can amplify deleveraging and impair market liquidity ([Brunnermeier and Pedersen, 2009](#); [Duffie et al., 2015](#); [King et al., 2023](#)), as can haircuts that tighten against falling prices ([Gorton and Metrick, 2012](#)). Repo-market evidence also emphasizes that haircut setting is institutionally segmented: in bilateral repo, haircuts reflect counterparty and relationship frictions, while in cleared repo, uniform CCP haircuts can shift borrower composition across cleared and OTC segments via adverse selection when CCP haircuts change relative to OTC ([Julliard et al., 2024](#); [Chebotarev, 2025](#)). Procyclicality and measurement of system-wide collateral demand feature prominently in policy and empirical work.²³ Regulation and governance therefore emphasize risk-based coverage and stability through anti-procyclicality tools such as buffers, stressed lookbacks, and floors²⁴; they also emphasize integrity of the price inputs used in rulebooks.²⁵ Three additional premises are empirically and institutionally relevant. First, margin rules are mechanically tied to price-based inputs (settlement marks and model-implied volatility/VaR/ES). Second, participants have at times traded strategically on these inputs with margin payments or calls as an explicit motive, including “banging the close” ([Onur and Reiffen, 2018](#)).²⁶ Third, near distress the distribution of collateral calls can become a first-order objective, shaping liquidation decisions, as highlighted by the 2022 UK gilt crisis, European energy derivatives in 2022, and the 1974 Caisse de Liquidation des Affaires en Marchandises (CLAM) episode ([Pinter, 2023](#); [Avalos et al., 2023](#); [Bignon and Vuillemeys, 2020](#)).²⁷

²²[Cont and Kokholm \(2014\)](#), [Loon and Zhong \(2014\)](#), [Duffie et al. \(2015\)](#), [Aldasoro et al. \(2023\)](#).

²³See, for instance, [Murphy et al. \(2014\)](#), [Murphy et al. \(2016\)](#), [European Systemic Risk Board \(2020\)](#), [Bank of England \(2021\)](#), [Aldasoro et al. \(2023\)](#).

²⁴The PFMI states: “A CCP should adopt initial margin models and parameters that are risk-based” ([Committee on Payments and Market Infrastructures and International Organization of Securities Commissions, 2012](#)); see also [Biais et al. \(2016\)](#); [Basel Committee on Banking Supervision \(2022\)](#); [European Union \(2013\)](#); [European Securities and Markets Authority \(2018\)](#); [European Systemic Risk Board \(2020\)](#); [Bank of England \(2021\)](#); [European Central Bank \(2023\)](#); [Basel Committee on Banking Supervision and Committee on Payments and Market Infrastructures and Board of the International Organization of Securities Commissions \(2022\)](#).

²⁵“The board of trade shall list on the contract market only contracts that are not readily susceptible to manipulation.” 17 CFR §38.200 (Core Principle 3). See [Commodity Futures Trading Commission \(2026a\)](#).

²⁶For a direct margin-motive episode, the CFTC found that Norman Eisler and First West Trading manipulated daily settlement prices of P-Tech options to inflate account value and thereby avoid or sharply reduce margin calls; recomputed settlement prices implied a much larger margin deficit ([U.S. Commodity Futures Trading Commission, 2001, 2004](#)). The resulting clearing failure triggered litigation by Klein & Co. Futures ([U.S. Department of Justice, Office of the Solicitor General, 2007](#)) and prompted a CFTC staff report on settlement-price governance and margin-risk controls ([U.S. Commodity Futures Trading Commission, Division of Trading and Markets, 2001](#)).

²⁷Recent stress episodes highlight that volatility-driven increases in margin and collateral demands can propagate liquidity stress across instruments and markets. For instance, in March 2020, higher margins and tighter effective leverage on Treasury futures plus repo relative-value positions contributed to rapid deleveraging pressure and strains in cash Treasury market liquidity ([Avalos and Sushko, 2023](#); [Aramonte et al., 2023](#); [Duffie, 2020](#); [Barth and Kahn, 2020](#); [Basel Committee on Banking Supervision and Committee on Payments and Market Infrastructures and Board of the International Organization of Securities Commissions, 2022](#)).