Margin Is Not a Footnote: Reg-T Feasibility as a First-Class Backtest Output

Charlie Yan*

August 2026

Abstract

Santa-Clara and Saretto (2009) showed that accounting for margin calls can turn the Sharpe ratio of an attractive short-option strategy negative. This paper is an implementation note on that result: a backtest of a margined strategy that reports a Sharpe ratio without a margin path is incomplete in a way that can invert its conclusion, and we show what reporting it takes. We document a short-premium index options book whose uncapped configuration backtests attractively on every performance statistic—and whose daily Reg-T requirement, reconstructed from the tape's own positions, *peaks at 151–154% of the account's net liquidation value*. The strategy is infeasible: it breaches its own capital on precisely the stress day that generates its risk premium. The naked short-call sleeve alone peaks at 108–111% of the whole account. Adding a far wing to that sleeve—an action costing roughly one-fifteenth of the equivalent put wing, because of smile asymmetry—halves the book's peak to 77% and converts an infeasible strategy into a deployable one at approximately zero cost to expected P&L. None of this is visible in any performance statistic, because margin peaks are joint events: entries cluster, holds stack, and per-position requirements widen simultaneously on the same stress day, so the peak-to-median ratio runs 6.7 to 16.1× across configurations while every performance statistic is a moment of a single daily series. We propose a minimum reporting standard for margined backtests—the dated requirement path, the peak's anatomy, an intrinsic-bound gap table on actual daily inventory, and return on deployed margin alongside return on capital—and we document three portable findings encountered in building it: a stale measurement window understated the true peak by half; iron-condor netting cuts the peak requirement 15% with identical positions and P&L, bounded by expiry alignment; and at small scale collateral is usually not the binding constraint (signal supply is), so margin optimization pays exactly at the scale-up where it stops being optional.

1 The statistic nobody reports

Consider two strategies with identical Sharpe ratios, identical maximum drawdowns, identical turnover, and identical per-year P&L. One of them can be run; the other cannot, because on a single

*Draft v0.3 (journal-length), August 2026. Margin figures come from our own Reg-T implementation, computed daily from the tape's actual positions (strikes, widths, credits, spot), hand-verified on sampled days against the published formulas, with two independent implementations agreeing to ∼1%. They are *our arithmetic, not the broker's*: validation against broker-computed requirements is pending and is flagged throughout. Requirements are reported as percentages of a reference net liquidation value; because requirements scale with position count, these ratios are scale-invariant and the normalization is a reference figure rather than deployed capital. Performance basis where quoted: full cross fills, ex-commission, marked to market daily on the full calendar, 2020-09 to 2026-05. Code, pre-registration documents and a data-availability statement: https://github.com/charlieyanhx/margin-first-class.

day in April 2025 its collateral requirement exceeds the account that is running it. No statistic in the standard performance toolkit distinguishes them. That gap is this paper's subject.

The omission is not a matter of degree. For an unlevered long-only strategy, capital consumption is trivially the position value, and reporting it adds nothing. For a margined short-option book it is a separate stochastic process, driven by different variables than P&L, peaking at different times, and capable of terminating the strategy independently of its returns. A backtest that omits it has not modelled the strategy; it has modelled the strategy's profits conditional on the strategy being permitted to continue—a condition it never checked.

We make three contributions. Section 2 presents a worked case in which the margin path inverts the deployment conclusion for a book that every performance statistic approves. Section 3 explains why performance statistics are structurally blind to this, through the joint-event mechanism, and quantifies the blindness. Section 6 proposes a reporting standard that costs a few lines of code against a position file, plus three findings from building it that we expect to generalize.

2 The case

The book comprises three sleeves of index option spreads—short put verticals, short call positions, and a delta-hedged relative-value sleeve—over 2020–2026, with entries clustering in volatility episodes and multi-day holding periods. Its backtested performance was acceptable on every statistic the standard toolkit produces, computed on honest fills and daily marking.

Reconstructing the daily Reg-T requirement from the tape's actual positions, with two independent implementations agreeing to ∼1% and anchored on hand-verified days:

Configuration (% of reference NLV)	Median	Mean	P90	**Peak (date)**
Put sleeve, winged	5.5	7.9	17.5	44.1% (2025-04-15)
Call sleeve, **naked (as designed)**	11.2	15.5	38	**108–111%** (2025-04-09)
Call sleeve, winged	2.2	3.5	7.8	35.4% (2025-04-08)
Book, naked-call configuration	22	—	—	**151–154%** (2025-04-09)
Book, winged configuration	11.5	—	—	**76.9%**

Table 1: Audited daily Reg-T requirement by configuration. The naked-call book breaches its own account; the winged book does not.

The breach day is not exotic. It is the April 2025 stress episode: roughly one hundred open short calls near the money after a +9.4% melt-up, at a hand-verified requirement per contract in the five figures. The strategy's premium is compensation for precisely this kind of day; the margin system's response to precisely this kind of day is what makes the strategy impossible to hold through it. That is not a coincidence but the structure of the problem: the event that justifies the premium and the event that triggers the collateral call are the same event.

One presentational recommendation follows directly. Plotting the margin path *against the equity curve*, rather than alone, makes legible the relationship between what a strategy has earned to date and what it must post today—a relationship that is unfavourable early in any strategy's life, since requirements scale with position count from day one while cumulative P&L starts at zero. We know of no performance statistic that substitutes for looking at the two series together.

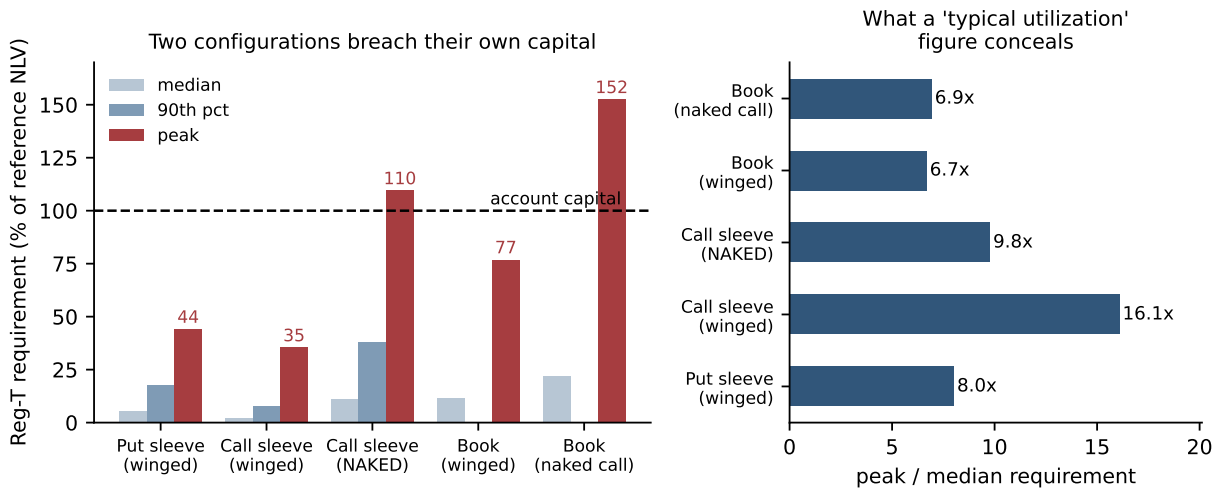

Figure 1: Exhibit 1. Left: the requirement distribution by configuration, against the account's capital line. Right: the peak-to-median ratio—what a "typical utilization" figure conceals.

3 Why performance statistics cannot see this

Margin peaks are *joint* events. On the same stress day three things happen at once: entries have clustered, because the signal fires repeatedly in episodes; holds have stacked, because positions from the preceding weeks are all still open; and each position's requirement has widened, because strikes have moved toward the money and per-position spread widths roughly doubled through the episode. The margin path is therefore the product of three correlated stress responses, while every performance statistic—mean, volatility, Sharpe ratio, even maximum drawdown—is a moment or an extremum of a *single* daily P&L series. The information is not compressed by those statistics; it is absent from their inputs.

The quantitative consequence is the peak-to-median ratio, which runs 6.7 to 16.1× across the configurations in Table 1 (Exhibit 1, right panel). Typical utilization is 11.5% of capital for the winged configuration and 22% for the naked one (16% median in the small-account variant): figures that, reported alone, are not merely incomplete but actively misleading, since each is roughly an order of magnitude below the number that determines whether the strategy can be run. Reporting the median without the peak is the collateral analogue of reporting a mean return without a drawdown.

Three portable specifics emerged from the audit.

(a) The stale-window trap. A previously recorded figure of "129% peak at production sizing" turned out to be the peak of a truncated early window—election week 2020—not exceeded again until 2022, after which position concurrency tripled. Measured on the full sample, the true peak was roughly double. Margin paths must span the full sample for exactly the reason performance must, and for an additional one: concurrency is a slow-moving structural variable, so an early-window margin measurement is not merely noisy but systematically biased low for any strategy whose position count grew.

(b) The cheap wing. Because of put-call smile asymmetry, the far call wing that converts the naked sleeve to defined risk costs roughly one-fifteenth of the equivalent put wing, while halving

the book's peak requirement—the cheapest unit of feasibility in the book, at approximately zero cost to expected P&L. This reframes what wings are for. In this book the put wing recovers only 22% of worst-case blowup losses, so as insurance it is poor value; as collateral relief it is decisive. *Wings are structure, not insurance*, and their margin function dominates their hedging function in the deployment decision.

(c) Netting. Pairing put and call spreads of matching expiry into iron condors— identical positions, identical P&L, since Reg-T charges the greater side rather than the sum—cuts the peak requirement 15% and the median 13%. The binding constraint is expiry alignment: the two sleeves choose expiries independently and coincide only 31% of the time. Deliberately aligning them should push the netting rate toward its limit at near-zero strategy cost, a one-variable change we have not yet tested. Broker recognition of the pairing must be confirmed before any of this is relied upon.

4 The gap table, and why history understates it

Realized drawdown history is the wrong basis for sizing a short-premium book whenever the sample lacks a generational stress day. Ours spans 2020-09 onward and contains neither March 2020 nor February 2018—a fact easy to state and easy to forget when a five-year backtest looks long. We therefore report, alongside the margin path, an *overnight gap table*: for every day's actual inventory, the position-level intrinsic-value bound on the overnight loss for gaps out to ±60%.

The construction is deliberately model-free. It asks what each open position is worth at expiry-equivalent intrinsic value under a hypothesised gap, sums across the actual inventory of that day, and takes the worst day. It requires no volatility model, no correlation assumption, and no distributional claim—which is what makes it a bound rather than a forecast. On the studied small-account configuration a −15% overnight gap bounds the loss at 49% of the account on the worst inventory day, and no gap on any day in the sample reaches forced liquidation: a statement made possible only by the defined-risk wings, and one that no realized-history statistic can make. Against that bound, the realized maximum drawdown understates the structural tail by roughly six times.

We regard the gap table as the natural companion to the margin path because the two answer the same question from opposite directions. The margin path asks: on the worst day the sample actually contained, could the strategy post collateral? The gap table asks: on a worse day the sample did not contain, would the strategy survive at all?

5 Return on margin versus return on capital

The book's return on margin actually deployed is roughly 111% per year. Its return on total capital is lower by exactly the reciprocal of utilisation—about six-fold at the 16% median of the small-account variant, placing it near 18% per year. The gap is not inefficiency. The idle capital is the reserve that makes the peak days survivable, and its opportunity cost is properly a cost of the strategy, not a shortfall in its execution.

Reporting both numbers prices the reserve explicitly. Reporting only return on capital hides the strategy's true capital intensity and makes it look worse than the trading is; reporting only return on deployed margin overstates scalability by pretending the reserve is free. The pair, together, is what a capital allocator needs, and the difference between them is the quantity a portfolio-margin regime or a netting improvement actually attacks.

One caution against over-optimizing here. At small scale, collateral is usually *not* the binding constraint in this book—median utilization is 16% in the small-account variant, and what binds is signal supply, roughly two entries per day. Margin optimization therefore pays at the scale-up where the peak becomes binding, which is exactly the point at which it stops being optional. Work on it before then and it is a solution without a problem; discover it after and the strategy has already been sized on a false premise.

6 A proposed reporting standard

For any margined backtest we suggest reporting, as first-class outputs alongside performance:

1. **The requirement path**: median, 90th percentile and peak, each dated, over the full sample.
2. **The peak's anatomy**: concurrency, per-position widening, and entry clustering on the peak day—the three multipliers, separated.
3. **A feasibility statement**: peak as a percentage of capital, with the margin regime named (Reg-T or portfolio margin) and the conclusion's sensitivity to that choice stated.
4. **An intrinsic-bound gap table** on actual daily inventory, not a distributional tail estimate.
5. **Return on deployed margin next to return on capital**, with the reserve priced.
6. **Where infeasible, the cheapest structural remedy** and its P&L cost—in our case, a wing costing approximately nothing that halved the peak.

Items 1–3 require a few lines of code against the position file a backtest already produces; none needs broker connectivity, though broker validation is the final word and ours is pending. We would go further and suggest that a margined backtest reporting a Sharpe ratio without at least item 1 should be regarded the way a backtest without transaction costs is now regarded: not wrong, exactly, but not yet a result.

7 The obvious alternative, tested: volatility-conditioned sizing

A reader who accepts that the naked configuration is infeasible will ask why the remedy should be a structural one—buying a wing—rather than a behavioural one: size down when volatility is high. The question deserves an answer rather than an assertion, because the behavioural fix is cheaper to implement and requires no premium outlay. We tested it.

The channel is real. Requirement scales with spread width, and width is set by delta-selected strikes, so a volatile tape mechanically produces wider spreads. On our put sleeve the rank correlation between the entry day's VIX percentile (prior close, trailing 252-day window, strictly causal) and the resulting spread width is $+0.665$, and entries in the top VIX quintile are $1.70\times$ wider than those in the bottom. A filter that declines high-volatility entries therefore has a genuine mechanism through which to cut the peak.

It is nonetheless dominated. We swept a causal rule—skip an entry when the prior close's VIX sits above its trailing q-quantile—and measured the change in peak requirement against the change in Line-3 P&L. The correct control is not a random gate but *uniform de-sizing*, which scales requirement and P&L by the same factor and therefore traces the line $\Delta\text{peak} = \Delta\text{P\&L}$ exactly. Any lever worth having must sit above it.

Skip entries above	Δ peak requirement	Δ P&L	versus uniform de-sizing
60th percentile	−60.8%	−73.0%	−12.2 pp worse
70th percentile	−55.5%	−65.4%	−10.0 pp worse
80th percentile	−17.6%	−52.4%	−34.8 pp worse
90th percentile	−10.1%	−30.6%	−20.4 pp worse

Table 2: Volatility-conditioned sizing against the size-matched control. At every threshold, shrinking every position uniformly to the same peak requirement retains more P&L than filtering on volatility does.

Why: it is one mechanism, not two. High-volatility entries are wider *and* carry the richest premium—the variance premium is concentrated exactly where the collateral is. Declining them sheds requirement and P&L together, and P&L marginally faster. The volatility signal is not identifying dispensable margin; it is identifying the book's best trades, which happen to be its most expensive to carry.

A methodological aside worth more than the result. Our first control was a duty-matched random gate—skip the same *number* of entries at random—which the rule passes at the two lower thresholds. That null is too weak here. When a lever's cost and its benefit are both proportional to position size, a count-matched control cannot discriminate, because it holds the wrong quantity fixed. The size-matched control reverses the verdict at every threshold. We report this because the count-matched gate is the more familiar null, and in this class of problem it will pass rules that do nothing.

The practical implication returns us to the wing. The structural fix is not merely *a* remedy for infeasibility; on this book it beats the obvious behavioural alternative, and it does so because it removes the tail of the requirement distribution without touching the entries that generate the return.

8 Relation to prior work

Margin and leverage constraints are central to modern asset pricing. Frazzini and Pedersen (2014) build the betting-against-beta factor on the premise that leverage constraints bind and are priced; Brunnermeier and Pedersen (2009) model the funding-liquidity spiral by which margin requirements and asset prices reinforce each other in stress—the macro version of the mechanism that makes our peak coincide with our premium event. Geanakoplos (2010) formalize collateral-constrained equilibrium; the margin-based asset-pricing literature that follows treats the requirement as a state variable rather than an afterthought.

Closest to our subject, Hitzemann et al. (2021) show that equity option returns carry a *margin premium* determined by an option's contribution to dealers' overall requirements—margin as a priced state variable. That is the mirror of our question: they ask what margin does to expected returns, we ask what it does to feasibility. **The closest antecedent, and it is very close.** Santa-Clara and Saretto (2009) document short option strategies with very high Sharpe ratios and then show that *accounting for margin calls turns the Sharpe ratio of some of the best strategies negative*, because margin requirements force liquidation precisely during sharp market declines. That is this paper's thesis, published seventeen years earlier: both the inversion and its mechanism—that the event justifying the premium and the event triggering the collateral call are the same event—are theirs, not ours. We say so plainly rather than rediscovering it. What this paper adds to their result

is operational rather than conceptual: a quantification of *why* performance statistics cannot see it (the joint-event decomposition, peak-to-median 6.7–16.1×), a concrete reporting standard, an intrinsic-bound gap table as an alternative to realized-drawdown history, and one specific empirical finding we have not seen elsewhere—that smile asymmetry makes the feasibility-restoring wing cost roughly one-fifteenth of its put-side equivalent while halving the book's peak requirement—and the test in §7 showing that the structural fix dominates the obvious behavioural one, volatility-conditioned sizing, at every threshold. A reader who knows Santa-Clara and Saretto (2009) should read this paper as an implementation note on their result, not as an independent discovery.

What is and is not novel here. We initially believed margin reporting to be absent from option-strategy backtesting; that overstates the case and we correct it. Retail and vendor backtest publishers routinely report capital utilization, and some size starting capital *to* the margin peak (targeting 80–100% maximum utilization), which is a stronger practice than merely reporting a median; return-on-margin alongside notional return is likewise established in that literature. Institutional risk systems compute requirements continuously. What we have not found, and what we claim, is narrower and we state it precisely: (i) a worked case in which the margin path *inverts* a deployment conclusion that every performance statistic approves; (ii) the joint-event decomposition explaining why performance statistics are structurally blind to it, quantified as a 6.7–16.1× peak-to-median ratio; (iii) the intrinsic-bound gap table on actual daily inventory as a complement to realized-drawdown history; and (iv) the specific empirical result that smile asymmetry makes the feasibility- restoring wing cost roughly one-fifteenth of its put-side equivalent. Readers who already size to peak utilization have item (i) covered; items (ii)–(iv) we believe are new.

9 Limitations

Our Reg-T implementation is hand-verified against published formulas and cross-checked between two independent builds, but it is not yet validated against broker-computed requirements; house minimums, add-ons and portfolio-margin models can all bind earlier than our arithmetic suggests, and the direction of that error is unfavourable. The portfolio-margin figures in the underlying audit are intrinsic-only approximations. The case is one book, one underlying, one margin regime, and a sample lacking a generational stress event—the last being precisely why Section 4 exists. The proposed standard, however, does not depend on the case's particulars: any strategy whose collateral requirement is a separate stochastic process from its P&L needs the requirement reported, and most margined strategies qualify.

References

Brunnermeier, M., Pedersen, L. (2009). Market liquidity and funding liquidity. *Review of Financial Studies* 22(6).

Frazzini, A., Pedersen, L. (2014). Betting against beta. *Journal of Financial Economics* 111(1).

Geanakoplos, J. (2010). The leverage cycle. *NBER Macroeconomics Annual* 24(1).

Hitzemann, S., Hofmann, M., Uhrig-Homburg, M., Wagner, C. (2021). Margin requirements and equity option returns. Working paper, SSRN 2789113.

Santa-Clara, P., Saretto, A. (2009). Option strategies: good deals and margin calls. *Journal of Financial Markets* 12(3).