Factor Model Analysis of Margin Backtesting for CCPs

Rama Kumanduri, Shahrzad Karimi

Model Risk Management
Options Clearing Corporation
April 19, 2026

Abstract

CCPs across the globe conduct backtesting of margin model performance for clearing member accounts to ensure model performance meets the desired 99% coverage requirement. The PFMI guidelines provide for aggregate coverage and analysis of backtesting exceptions to guide model enhancement decisions. CCPs typically employ traditional methods such as Kupiec or Christofferson tests, along with distributional analysis such as Probability Integral Transform tests to probe model performance. Some CCPs also impose additional margin charges for backtesting under-performance based on tests of margin adequacy.

Regulatory guidance also exists for core model backtesting on representative portfolios, but no guidance exists on linking backtesting performance on unit portfolios of representative risk factors to aggregate model performance. We provide an ex-ante margin exception analysis, i.e., a method to estimate expected number of margin backtesting exceptions using a factor model that permits model developers and risk managers to test model performance on a fixed portfolio. Margin model backtesting exceptions can be decomposed into factors connecting to unit portfolios, and deviations from expectations regarding model performance can lead to insights on model deficiencies, including modeling of risk factor volatilities and correlations. Additional scenario analysis for stress events can guide decisions on margin procyclicality mitigation based on expectations of model performance targeted at 99% confidence[1].

Keywords: margin backtesting, factor model analysis, procyclicality control

Contents

[1]This research was supplemented with analysis from Generative AI tools, Claude and ChatGPT

1 Introduction

Central Counterparties (CCPs) across the globe are increasingly using Value-at-Risk (VaR) or Expected Shortfall (ES) based models to calculate initial margin requirements (Boudiaf et al. (2023)). These risk-based modeling approaches lead to capital efficiencies for clearing members and participants, but require model performance monitoring, validation and continued enhancements to ensure the models continue to meet supervisory standards as market conditions or portfolio compositions change.

The Principles of Financial Markets Infrastructures (PFMI), (CPMI (2014)) provide guidance to CCPs on margin calculations and backtesting performance monitoring involving comparing the historical profit and loss (P&L) against the initial margin (IM) collected on clearing member's margin portfolios. An exception or exceedance occurs when the P&L (measured as a loss) exceeds the deposited IM.[2] The PFMI guidance regarding backtesting exceedances states that:

> In case backtesting indicates that the model did not perform as expected (that is, the model did not identify the appropriate amount of initial margin necessary to achieve the intended coverage), a CCP should have clear procedures for recalibrating its margining system, such as by making adjustments to parameters and sampling periods. In addition, a CCP should evaluate the source of backtesting exceedances to determine if a fundamental change to the margin methodology is warranted or if only the recalibration of current parameters is necessary.

While the PFMI does not provide a definition of coverage, it is commonly interpreted to be aggregate coverage over a period, i.e., 100% - the exception rate. The latter is defined the ratio of the total observed exceptions to the total observations. The public quantitative disclosure standards for central counterparties require CCPs to disclose the overall coverage rate, including the number of exceptions, the largest deficiency, the median deficiency and the cause of the deficiency (where known). CCPs are not required to disclose the number of accounts that do not meet the required confidence.

There is less guidance for CCPs on how to conduct model backtesting and the count of the backtesting exception rate due to changes in portfolio composition. It is expected that each account should satisfy the desired confidence level, and if the exception rate is higher than the target, then steps that should be taken to adjust margin model parameters or sampling period, and/or undertake a replacement of the margin model. CCPs must continue to provide clearing services

[2] Initial margin should meet an established one-tailed confidence level of at least 99 percent with respect to the estimated distribution of future exposure.

for all products and replacement of models is a large undertaking. It is also possible that specific market conditions and/or idiosyncratic risk factor behavior caused the margin exceedances, and adjusting parameters or replacing a model will be a backward-looking approach to deal with past failures rather than a forward-looking approach to address potential future exceptions. Some CCPs do apply an anticipatory margin charge ahead of events, though it is not stated if the charge is driven by model backtesting performance. [3]

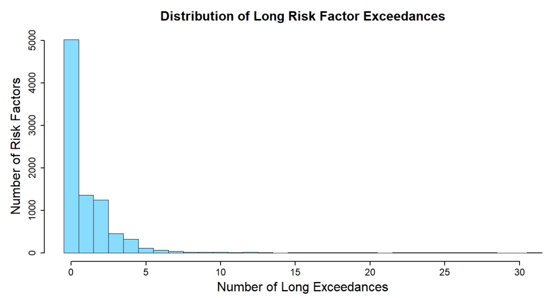
The European Securities Markets Authority (ESMA) published an opinion paper in 2023 (ESMA (2023)) regarding CCP backtesting and distinguished between **Core Model Backtesting** and **Margin Adequacy Backtesting**. As stated in §20 of the ESMA opinion, margin adequacy back testing evaluates whether the CCP had sufficient initial margin to cover its losses on a specific historical day assuming a clearing member defaults. This type of backtesting uses actual portfolios, and includes transactions between the time of last margin collection and the assumed time of default. Many CCPs (DTCC, OCC, LCH) have implemented an ex-post backtesting charge for accounts that do not meet the required confidence over the lookback period. One of the drawbacks of such ex-post charges is that it may not reflect the current portfolio composition, particularly changes to idiosyncratic risk factor categorization. A more serious issue with backtesting charges based on 3 exceedances in one year is the increase in procyclicality of margin collected. If a market stress event causes an account to experience a third deficiency, an automatic backtesting charge is imposed, leading to an increase in margins at the same time the CCP is managing the inherent procyclicality of the margin model. The charge proposed at LCH is different, as it is based on type-II model errors (LCH (2023)). The method involves backtesting with additional hypothetical scenarios, but not necessarily unit portfolio exceedances.

Regarding core model backtesting, ESMA clarified the expectations of CCP backtesting under EMIR to require unit backtesting in addition to portfolio-wide assessment. The coverage requirements apply to both portfolio and representative unit portfolios. The guidance covers unit portfolios consisting of single positions (long/short), along with the combinations such as off-setting spreads, calendar spreads, curve trades and typical option strategies for CCPs clearing options. ESMA notes that the selection process for the representative subset must be free of any bias, such as excluding a certain part of the yield curve or avoiding particular option strikes. BIS guidance on Supervisory stress testing for CCPs (BIS (2018)) recommends working off a small set of representative core risk factors to gauge system-wide effects.

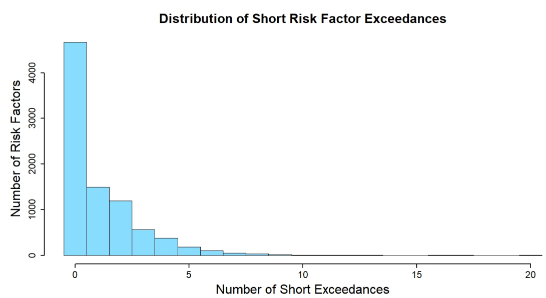
However, the ESMA guidance is not prescriptive on the selection of the representative risk factors and the backtesting exceptions for unit portfolios, other than that the exceptions should not be clustered. CCP portfolio margin must include all the risk factors in the market they support, and it is quite likely that some unit portfolios do not satisfy the required coverage. Figure 1 shows the exceedance count for single position portfolios of several thousand US Equities and ETFs using a historical volatility calibration model. Clearly, there are several hundred risk factors with more than 3 exceedances in the calendar year (coverage below 99% with 252 business days). While these may not be representative of portfolio exposures in a given year, potential changes to markets and trading activities can lead to significant growth in portfolio exposure to risk factors with under-coverage.

It is well known that single name securities have large price movements around earnings events. The options market responds to this with increased volume and changes to implied volatility to anticipate the expected price movement, as recently studied by Alexiou et al. (2025). The increased

[3] DTCC applies an ex-ante 10% volatility move charge ahead of certain known market events. "https://www.dtcc.com/-/media/Files/pdf/2026/3/23/GOV2139-26---Volatility-Event-Charge-Events.pdf".



(a) Long Exceedances for US Equities



(b) Short Exceedances for US Equities

Figure 1: Exceedances for unit portfolios of US Equities: Jan-Dec 2024

volume and price movement causes backtesting exceedances at unit portfolios and account level. Our analysis of exceedances for the top 50 securities in the S&P 500 index shows (Table 1) that some risk factors do not satisfy the 99% coverage, even with conservative margin models. The bulk of the exceedances ($> 90\%$) are in the five days surrounding corporate earnings announcements. As earnings reporting is quarterly, some risk factors have 4 large price movements in a year, making it difficult for a historical simulation model with volatility scaling to capture these price changes in margin model scenarios. Such event risk is also prevalent in interest rate and commodities markets.

Margin Coverage	Number of Risk Factors (Long)	Number of Risk Factors (Short)
100.00%	17	17
99.80%	11	13
99.60%	11	6
99.40%	6	7
99.20%	3	3
99.00%	2	1
98.80%	-	2
98.60%	-	1

Table 1: Long/Short coverage rates for Top 50 names in S&P 500

Existing literature on model performance backtesting does not integrate unit backtest performance to account level performance. The traditional tests such as Kupiec and Christoffersen tests are focused on clustering of exceedances, and designed for banking institutions, as an excessive number of VaR exceptions leads to an increase in the capital multiplier. The increase in regulatory capital acts as a mitigant for model performance deficiencies at banking institutions, but CCPs are not required to increase margin requirements due to backtesting exceedances, nor is it always desirable to do so. It is not straightforward to extend whole distribution or probability integral transform based tests from unit portfolios to large clearing member portfolios.

The marginable positions at CCPs are typically balanced as the long and short positions must net out, but portfolio exposures do not need to be balanced. An exceedance in a unit portfolio of a representative risk factor does not lead to deficiencies at all accounts, as some members will experience a gain while others have losses. Only accounts with concentrated exposure to the underperforming unit portfolio are likely to have exceedances.

The Filtered Historical Simulation (FHS) method for initial margin calculations at CCPs is studied in Gurrola-Perez and Murphy (2015). Their study also noted that FHS models don't handle correlation breakdowns very well. A concentrated spread position in energy futures led to a large default at Nasdaq Clearing on September 10, 1998. The empirical and theoretical analysis by Li et al. (2024) shows that there was a scenario with stress returns on correlation breakdown on August 1, 2018, but the filtered VaR model does not have a significant adjustment compared to the unfiltered variants.

Extensions of FHS methods to use factor models were recently proposed by Du and Nesmith (2025). They include latent factors from PCA analysis to improve the performance of FHS. They note that *"These results indicate the challenge of producing a margin system that works for any portfolio, and illustrates how CCPs need to closely monitor margin performance, particularly for smaller and idiosyncratic portfolios"*. The authors also note that specific portfolios might bring less systematic risk, but could still result in large losses to a CCP especially when the exposure is sizable and concentrated.

We propose a factor model analysis to link unit portfolio exceedances to margin account exceedances. The use of a factor model is a standard approach in portfolio analysis. While margin models are typically based on Historical Simulation and variants (FHS) and/or Monte Carlo simulation of all the risk factors in the portfolios, stress testing approaches use a parsimonious factor model representation to probe the portfolio exposures to systematic risk factors. Such models may also include an idiosyncratic residual component. It is possible for portfolio concentrations to accumulate for risk factors with low factor loadings to systematic risk factors. We use the margin portfolio concentration analysis to link the unit portfolio exceedances to the aggregate CCP coverage.

In our approach, the representative unit portfolios become the factors in the analysis. These are applied to a fixed position portfolio, where the market data inputs are changing over the backtesting approach. The factor loadings can be used to detect the concentration of a clearing member account in the unit portfolios. The severity of the loss based on the factor-based P&L can be used to devise the targeted add-on for the specific account. In Table 9 we list the potential backtesting issue along with recommendations regarding targeted add-on vs model re-calibration or re-development.

The PFMI guidance provides for additional margin charges for potential correlation changes in stress periods, and for idiosyncratic or concentrated exposures. These add-ons are difficult to calibrate at account level due to the rarity of account exceedances attributed to these causes (as aggregate coverage is generally well above 99%). These charges are also applied to all accounts irrespective of exposure. Our proposed approach provides a way to assess the conservativeness of the add-ons and to target specific accounts for additional margin.

2 Methodology and Case Studies

2.1 Notation and Inputs

The set of risk factors underlying the products cleared at a CCP will be denoted by $\mathcal{S}$. These may be equity and commodity price factors, interest rates, and credit spreads. The price of a risk factor i in $\mathcal{S}$ will be denoted by P_i.

Let $\mathcal{IM}$ denote the initial margin model used by the CCP. Any portfolio Π will have a margin $IM_t(\Pi)$ on day t, or IM when the portfolio is implicit.

The risk factor returns $r_{i,t}$ on day t are defined in a way that directly maps to the way the portfolio Profit & Loss (P&L) is calculated. The appropriate return definition depends on the asset class:

- **Equities/Commodities** Simple price change $\Delta P = P_t - P_{t-1}$, or log-return $r = \ln(P_t/P_{t-1})$.
- **Interest Rate Products** Absolute rate or yield changes $\Delta y_t = y_t - y_{t-1}$.
- **FX** Percentage or absolute change in spot rate.
- **Credit (CDS or bond spreads)**: absolute spread change.

For our analysis, the portfolio P&L can be computed with risk factor sensitivities (deltas, dv01, cs01 etc) or with full revaluation of the portfolio to account for convexity and other non-linear features.

The set of portfolio risk factors are denoted by $\mathcal{F}$ with individual factor prices $f_{k,t}$. These factors may be a combination of the above, such as spreads. Option portfolios will have underlying prices, implied volatility at different tenors, interest rates at key tenors, volatility skew etc as potential factors. Unit portfolios of options can be used to isolate the specific risk factor. For instance, a box strategy is primarily interest rate driven, whereas a delta-hedged option (with re-balancing) is exposed to implied volatility as the primary risk factor.

2.2 Factor Model Methodology

The portfolio for account a is fixed on day T, with historical market data price changes of the risk factors for N days prior to T used to calculate the vector of portfolio P&Ls, $\Delta V(\Pi)_t = V(\Pi)_t - V(\Pi)_{t-1}$ where V is the valuation function. The vector of P&L data is calculated with a fixed portfolio to isolate the market data changes.

The model specification is the standard multi-factor framework with K hypothetical factors.

$$\Delta V_t(a) = \sum_{k=1}^{K} \beta_{a,k} f_{k,t} + \epsilon_t \tag{1}$$

In general, diversified portfolios are expected to have high factor betas with respect to the systematic drivers (such as SPX index). When the focus is on spread portfolios, it is better the factors are the individual spreads of interest to ensure the significant betas are detected. This is achieved by running the one-factor regression to isolate the effect of holding other factors constant.

The betas in the regression are directional, and related to the position contribution of the respective unit portfolio. This ensures that a relationship can be found between the exceedances of the unit portfolio and the exceedances of the portfolio. The sign of beta, $\beta_k > 0$, means the account is effectively long factor k (loses when factor falls); $\beta_k < 0$ means effectively short (loses when factor rises).

An account-level exceedance is defined by the indicator $I_t = 1$ if P&L $\Delta V_t < -IM_t$. The convention is for the IM to be represented as a positive quantity. For portfolio P&L, only the loss tail is relevant.

Unit portfolio exceedances for factor f_k are directional.

- Define $J_k^+(t) = I(f_{k,t} < -IM_k^+(t))$ for a long portfolio, and
- Define $J_k^-(t) = I(f_{k,t} > IM_k^-(t))$ for a short position.

where $IM_k^{\pm}(t)$ is the directional VaR measure for the unit portfolio. The different directional IM for long/short portfolio reflects the **Tail Asymmetry** for the factor, and leads to different exceedance rates.

Given the N days in the backtesting period, define the unconditional exceedance rates for the accounts and unit portfolios as

- $\hat{p} = \frac{1}{N}\sum_{t=1}^{N} I_t$
- $\hat{q}^+ = \frac{1}{N}\sum_{t=1}^{N} J_k^+(t)$

Our goal is to provide a linkage between $\hat{p}$ and $\hat{q}^+$, $\hat{q}^-$ for attribution of account exceedances to factors.

We also define **Margin Shortfall** to be the loss beyond the margin.

$$S_(t) = I_t \times (V_t - IM_t) \tag{2}$$

The severity of the expected margin shortfalls can be used to determine the additional margin charged to an account.

2.3 Factor Conditional Analysis

The factor-conditional account exceedance rate measures how frequently an account breaches on days when factor k itself has a breach.

$$\hat{p}_{a|k} = \frac{\sum_{t=1}^{N} I_{a,t} J_{k,t}}{\sum J_{k,t}} \tag{3}$$

The unconditional probability of account exceedances is the exceedance rate, i.e., $\hat{p}_a = \hat{p}$. Each factor has a corresponding sign, based on the direction of the beta. For diversified portfolios under a well calibrated model, if account exceedances and factor exceedances were independent, then the conditional probability is small.

For each account and factor, we define a **Concentration Ratio** as:

$$C_{a,k} = \frac{\hat{p}_{a|k}}{\hat{p}_a} \tag{4}$$

and a conditional coverage as:

$$FCR_k = 1 - \hat{p}_{a|k} \tag{5}$$

The factor contribution to the magnitude of margin exceedance can be quantified by the factor model betas.

$$\phi_{a,k} = \sum |\beta_{a,k} f_{k,t}| / \sum |S_{a,t}| \tag{6}$$

where the sum in the denominator is only on the days when there is an account exceedance.

The expected number of exceedances attributable to the factor is derived from the definition of the concentration factor:

$$E[\text{Number of account exceedances from factor k}] \approx 252 \times C_{a,k} \times \hat{p} \times \hat{q}_k \tag{7}$$

2.3.1 Illustration 1

Consider a case where an account has a large exposure to a single name security. If the unit portfolios in the security have exceedances near earnings date, then a high concentration in the security signals the likelihood of the account level exceedance. Table 2 shows hypothetical scenarios for two years of data. If the account shows 8 exceedances, and if most of these are attributed to the risk factor, the concentration factors will be large, signaling a potential breach. The account exceedances already demonstrate that the margin model does not satisfy the required 99% confidence, and it would be prudent to include a margin add-on for the account ahead of earnings announcement. As this is a known event risk, the additional margin is targeted for the event days rather than applied on a daily basis.

Unit Exceedances	Account Exceedances	Joint Exceedances	$p(a\|k)$	$p(a)$	$C(a,k)$
5	1	1	0.20	0.002	100.00
8	6	3	0.375	0.012	31.25
8	8	8	1.00	0.016	62.50

Table 2: Sample Data for Portfolio with Single Name Securities Exposure

2.3.2 Illustration 2

As a simple example, suppose that a portfolio margin model for option portfolios at CCP is deficient in capture of interest rate risk changes in the scenarios. In the Black-Scholes pricing framework, the sensitivity of the option portfolio to interest rate risk is called the Rho. Unit portfolios can be constructed with primary exposure to interest rate risk. These are the well known box spreads, used for funding transactions by market participants[4]. In this case, the account IM will be relatively small due to lack of interest rate scenarios in the model.

Assume a new portfolio enters the margining system with exposure to interest rates via box spreads. The linear regression betas in Equation 1 will correspond to the Rho sensitivities via the discount factors. In this case, the bulk of the exceedances are attributed to interest rate risk, as the portfolio margin is small and interest rate is the only risk factor. The concentration factor will be large, triggering alerts regarding under-margining. A margin add-on can be imposed based on the severity of the shortfalls in Equation 2. The sample data in Table 3 provides indicative numbers for the concentration ratio. In this example, the relationship between the unit portfolio and account level exceedances is detected to demonstrate the joint relationship. It is possible that in some cases, the account will have a larger number of exceedances than the unit portfolio, diminishing the concentration effect. In such cases, the factor contribution to the account exceedance magnitude can be measured via the beta coefficients.

Unit Exceedances	Account Exceedances	Joint Exceedances	$p(a\|k)$	$p(a)$	$C(a,k)$
5	1	1	0.20	0.002	100.00
8	4	3	0.375	0.008	46.88
12	6	6	0.50	0.012	41.67
12	6	6	0.50	0.012	41.67
15	10	8	0.533	0.02	26.67

Table 3: Sample Data for a hypothetical portfolio of Box spreads

2.4 Additional Factor Analysis Ratios

We define additional factor analysis ratios for directional and aggregate exceedance rate calculation.

[4] See `https://www.cmegroup.com/articles/2024/index-options-box-spreads-as-financing-tool.html`

Directional Concentration Ratio This ratio measures the probability of the breach based on the direction of the unit portfolio, for example, the relevant direction for the long portfolio is a decrease in the factor.

$$\hat{p}^{d}_{a|k} = \frac{\sum_{t=1}^{N} I_{a,t} J^{d}_{k,t}}{\sum J^{d}_{k,t}} \tag{8}$$

where the superscript d is either long or short for the factor.

The directional concentration ratio defined as:

$$C^{d}_{a,k} = \hat{p}^{d}_{a|k} / \hat{p}_{a} \tag{9}$$

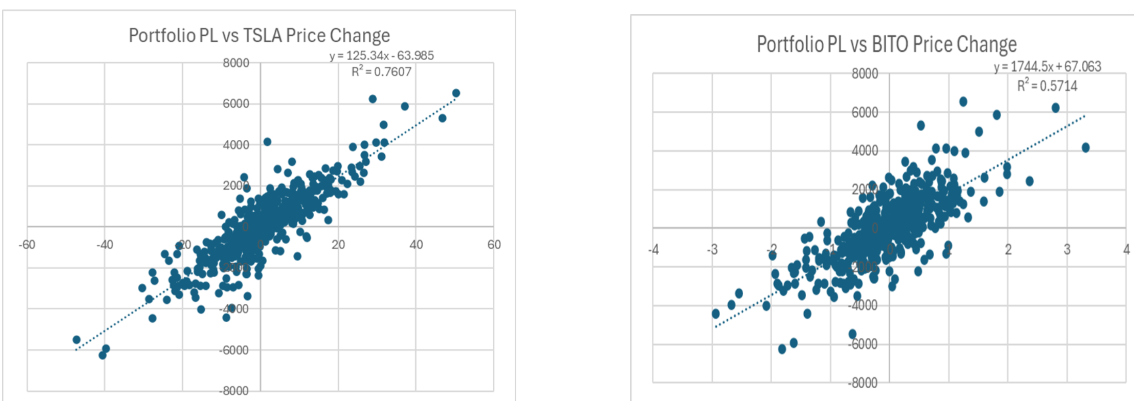
Directional Asymmetry Ratio The asymmetry ratio compares the conditional exceedance rate on adverse factor days versus favorable factor days

$$A_{a,k} = \hat{p}^{d}_{a|k} / \hat{p}^{d'}_{a|k} \tag{10}$$

as the ratio of the opposite direction conditional probabilities. Here d' is the opposite direction to d.

2.5 Case Study 1: Equity Portfolios

We study a basket of two equity stocks to measure the contribution of the idiosyncratic risk factors. The portfolio is a basket of two stocks, 1200 shares of BITO and 100 shares of TSLA. The single factor regression of the portfolio PL shows the betas to be 1744 and 125, reflecting the contribution of the correlation of the risk factors. The data in Figure 2 shows that the regression has a high R^2 and the betas correspond to the position contribution.



(a) Portfolio PL vs TSLA price change

(b) Portfolio PL vs BITO price change

Figure 2: Regression Betas

The portfolio concentration factor in Table 4 shows high concentration in both BITO and TSLA due to number of unit exceedances, but the portfolio P&L attribution in Figure 3 shows that the sensitivity is primarily to the TSLA risk factor.

Portfolio	Exceedance	Coverage	C(a,k)
BITO	7	98.60%	25.90
TSLA	11	97.80%	20.40
Account	7	98.60%	-

Table 4: Concentration Factor for Basket of TSLA and BITO

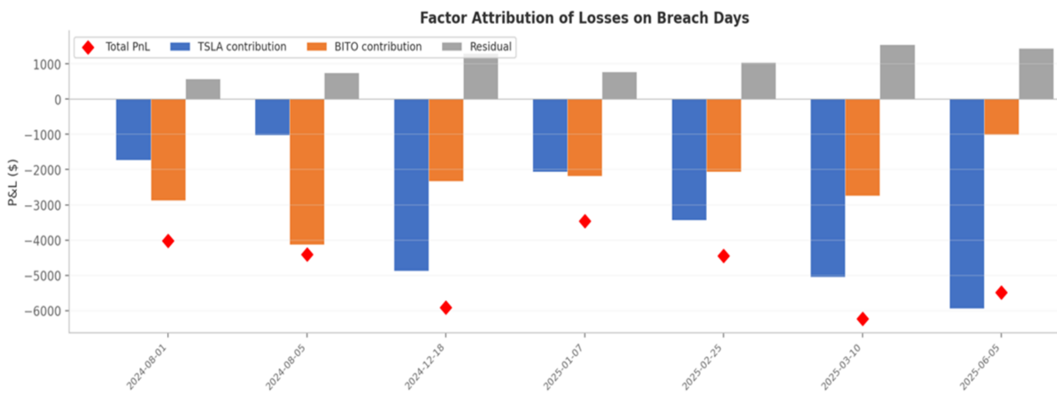


Figure 3: Risk Factor Contribution to Portfolio P&L

In this case, the concentration factor to TSLA and the factor P&L contribution make the account more likely to have margin exceedances, particularly near earnings. A targeted margin add-on ahead of earnings ensures the total initial margin is more likely to meet the expected confidence.

2.6 Case Study 2: Commodity Portfolios

2.6.1 WTI Spread

We study a spread portfolio constructed from West Texas Intermediate crude oil futures, consisting of a long position of 1000 contracts in the front-month future (WTI1) and a short position of -1100 contracts in the second-month future (WTI2). Unlike Case Study 1 where the equity portfolio is exposed to two outright asset movements, this account is driven by the relative dynamics of nearby contracts along the WTI term structure, and therefore the account risk is shaped both by the individual contract sensitivities and by the interaction between the two legs.

Figure 4 illustrates the historical trend of account P&L against the corresponding VaR at 99% confidence level. The figure indicates that prior to March 2026, the account coverage was above 99%, suggesting the margin model performed adequately over most of the sample period. However, the majority of account exceedances occurred in March 2026. This concentration of exceedances is associated with geopolitical developments during that period, which affected the oil market and led to increased stress in the nearby part of the curve. As a result, the model underestimated the risk of the account during this short stress window, and the account became under-margined.

Figure 4: Historical Account P&L and VaR, WTI Spread

The portfolio concentration factor in Table 5 shows that the account exhibits exposure to both WTI1 and WTI2, with materially higher concentration in WTI2. This indicates that the second-month contract plays a dominant role in driving portfolio risk, particularly under stressed curve conditions.

Additionally, factor conditional coverage suggests stronger alignment between unit-level exceedances and account-level breaches for WTI2. The factor-conditional coverage (the share of unit-level exceedances that coincide with account-level exceedances) is significantly lower for WTI2 than WTI1, consistent with the higher concentration in WTI2 and its greater influence during stress periods. This implies that the account-level tail events are driven by both spread dynamics and WTI2.

Portfolio	Exceedance	Coverage	C(a,k)	Factor-Conditional Coverage
WTI 1	8	98.4%	5.7	87.5%
WTI 2	11	97.8%	16.5	63.6%
Account	11	97.8%	-	-

Table 5: Concentration Factor for WTI Long Spread Portfolio

Finally, Figure 5 summarizes the diagnostic interpretation. WTI1 falls in to the "diversification absorb shortfall region", indicating that its standalone risk is largely mitigated at the portfolio level. In contrast, WTI2 falls into "factor k under-margined", reflecting the combination of high concentration and weak conditional coverage. This provided clear evidence that the margin model underestimates the standalone and marginal risk contribution of WTI2. Consequently, a targeted margin add-on for WTI2 is justified to improve coverage during periods of curve stress.

2.6.2 Gold/Silver Spread

We analyze a spread portfolio consisting of a long position of 100 front-month gold future contracts and a short position of 2000 front-month silver future contracts. Similar to the WTI spread case, the account is exposed to the relative movement between two commodity factors rather than outright directional risk. As a result, portfolio risk depends on both individual factor behavior and the

	$C(a,k) \approx 1$ **(No Concentration)**	$C(a,k) \gg 1$ **(Concentrated)**
$q_k \approx 1-\alpha$ (Unit PASS)	**Model OK for factor k**	**Correlation / Offset Failure**
$q_k \gg 1-\alpha$ (Unit FAIL)	**Diversification absorbs shortfall** **WTI1**	**Factor k under-margined** **WTI2**

Figure 5: Diagnostic decision matrix based on concentration and unit backtest performance.

interaction between the two legs.

Figure 6 presents the historical evolution of account P&L relative to the 99% VaR threshold. The account exhibits 22 exceedances, corresponding to a coverage level of 95.6%, indicating a clear breakdown of the margin model over the sample period. Both underlying factors also fail their unit-level backtests, with gold exhibiting 28 exceedances (94.4% coverage) and silver exhibiting 30 exceedances (94 % coverage). This suggests that, at the standalone level, both factors are under-margined.

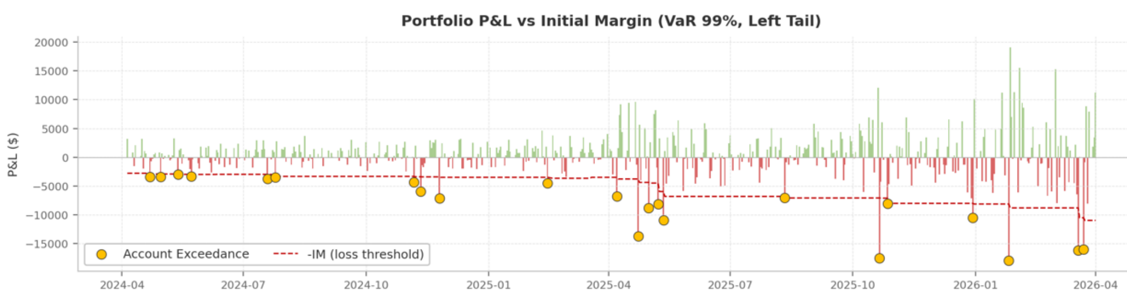


Figure 6: Historical Account P&L and VaR, Gold/Silver

Despite this, the alignment between unit-level and account level breaches differs materially across the two factors. In particular, the timing of silver exceedances shows little overlap with account exceedances, whereas gold exceedances exhibit stronger alignment with account stress event, see Figure 7. This distinction is further supported by the factor conditional coverage: silver has a relatively high conditional coverage of 93.3%, while gold exhibits a significantly lower conditional coverage of 39.3%. This indicates that gold-driven stress events are far more likely to propagate to the account level.

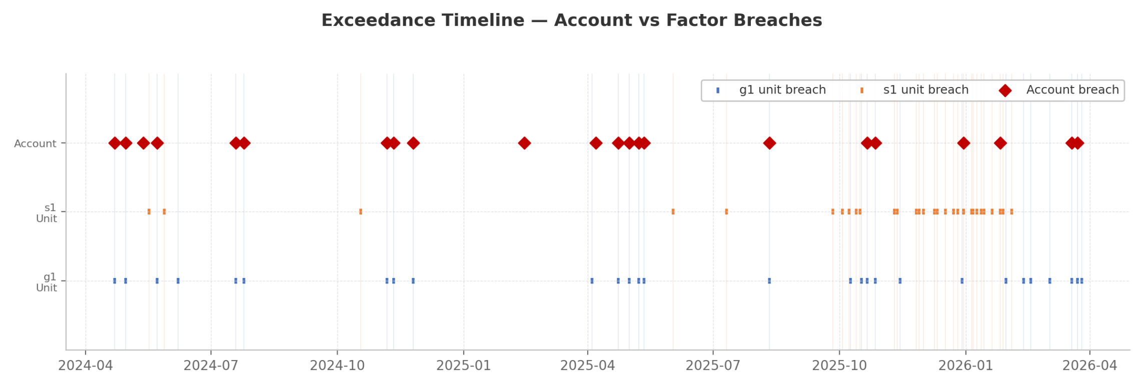


Figure 7: Timeline of account breaches and unit portfolio breaches.

The portfolio concentration factors reported in Table 6 provide additional insight. The account exhibits a high concentration in gold while the concentration in silver is very low. This implies, although silver has a higher number of standalone breaches, its contribution to account-level risk is limited due to its low effective exposure within the portfolio.

Portfolio	Exceedance	Coverage	C(a,k)	Factor-Conditional Coverage
Front Month Gold	28	94.4%	13.8	39.3%
Front month Silver	30	94.0%	1.52	93.3%
Account	22	95.6%	-	-

Table 6: Concentration Factor for Front-month Gold-Silver Spread Portfolio

Figure 8 summarizes the diagnostic interpretation. Based on the diagnostic matrix, the two factors fall into distinct regions. Silver is classified in the ”diversification absorb shortfall” category: although it fails the unit-level test, its low concentration and weak alignment with account breaches indicate that its shortfall does not materially propagate to the portfolio. In contrast, gold lies in the ”factor k under-margined” region, reflecting the combination of high concentration and strong linkage to account-level breaches.

This case highlights an important distinction between standalone model performance and portfolio relevance. A factor may exhibit poor standalone coverage without posing a material risk to the account if its exposure is sufficiently small or diversified. Conversely, factors with high concentration and strong exceedance alignment can drive portfolio-level model failure even if their standalone performance is only moderately worse. Accordingly, the results suggest that remediation efforts should prioritize improving the margin treatment of gold, while silver does not warrant a targeted margin add-on despite its higher number of unit exceedances.

	$C(a,k) \approx 1$ (No Concentration)	$C(a,k) \gg 1$ (Concentrated)
$q_k \approx 1 - \alpha$ (Unit PASS)	**Model OK for factor k**	**Correlation / Offset Failure**
$q_k \gg 1 - \alpha$ (Unit FAIL)	**Diversification absorbs shortfall** **Silver**	**Factor k under-margined** **Gold**

Figure 8: Diagnostic decision matrix based on concentration and unit backtest performance.

3 Literature Review

We provide a brief summary of recent literature on margin backtesting. The focus of the papers is on ex-post analysis of realized exceptions, whereas the ex-ante analysis is generally restricted to supervisory guidance.

The techniques for assessing the performance of Value-at-Risk models were pioneered by Kupiec (1995) and Christoffersen (1998). These tests implicitly assume a fixed portfolio for backtesting, due to the assumption that the sequence of exceptions comes from a stationary data-generating process, which is most consistent with an effectively fixed portfolio. The Kupiec test's IID requirement implies the exceedance probability on any day is constant, but this is violated for exceedances with known events (earnings, economic indicators ..). The Conditional Coverage Independence (CCI) test by Christoffersen evaluates whether breaches are temporally clustered rather than independently distributed as would be expected for a model with equal probability of an exceedance on a given day. It does not extend to clustering of exceedances in unit portfolios.

Gurrola-Perez (2018) validates backtesting for Filtered Historical Simulation models and demonstrates that traditional VaR backtests lack discriminatory power and proposes Asymmetric Scoring Rules via the Asymmetric Piecewise Linear (APL) loss function as a superior alternative. This aligns with the broader literature on strictly consistent scoring functions for quantiles that take the amount of the margin exceedance into account and is superior to the traditional approaches that do not consider the severity of the backtesting exceedance, and consistent with the approach taken by Murphy, discussed next.

Murphy (2023) provides for whole-distribution tests for testing initial margin models based on their predictions for the distribution of risk factor returns. It is a powerful approach to identify if the model under/over-estimates the tails. These tests use cumulative density function of the predicted portfolio P/L and compare to the realized portfolio P/L. As they take all the portfolio P/L realizations into account, they have more information and predictive power compared to the traditional Kupiec or Christoffersen tests. Similar methods are applied to backtesting of Value-at-Risk Models for banks in Chapter 4 of Lynch et al. (2023).

Hurlin and Perignon (2011) provide several ex-post margin backtesting tests focusing on realized exceedances measured by portfolio P&L , with tests based on the number, frequency, magnitude and timing of exceedances.

The book Lynch et al. (2023) provides the first unified framework for validating risk management models. The book discusses current practices and pitfalls that model risk users need to be aware of and identifies areas where validation can be advanced in the future. The chapters cover the performance of banks risk management models during the COVID crisis. While there are no data or case-studies presented for CCPs, the techniques can be directly applied to the models used by CCPs to evaluate performance.

The paper by Engle and Manganelli (2004) proposed Conditional Autoregressive Value at Risk (CAViaR) model that directly models the quantile behavior rather than estimating the entire distribution of returns and introduced a new Dynamic Quantile (DQ) test to evaluate model adequacy, checking if the probability of exceeding VaR is independent of past information.

The paper by Dumitrescu et al. (2012) introduces a new approach for backtesting Value-at-Risk (VaR) models using dynamic binary response models. The authors address key limitations of existing backtesting methods, particularly the popular Dynamic Quantile (DQ) test by Engle and Manganelli (2004). The analysis requires using a Maximum Likelihood Estimator, but similar tests can be adapted to factor model based backtesting.

4 Conclusion and Further Work

We present a new approach to probe CCP margin model performance to link unit portfolio backtesting performance to account level backtesting analysis using a factor model approach. The model backtesting is conducted on a fixed portfolio, ensuring changes to portfolio do not impact model performance results. This study is motivated by a desire to provide model validators, model developers and risk managers with tools to assess model performance on an ex-ante basis. An ex-post margin charge can be applied based on severity analysis of anticipated backtesting breaches (earnings events, FOMC announcements, economic indicator reporting such as GDP, unemployment figures etc.).

Unit portfolio backtesting (long/short) and coverage rates can be integrated into the account level margin coverage by factor beta and concentration analysis. This approach allows model developers, validators and risk managers to assess conditions under which the unit portfolio contribution can deteriorate and lead to aggregate coverages falling below target coverage. Such analysis can guide timeline for model parameter and calibration changes.

The preliminary results show that factor analysis for linear portfolios of Futures, Swaps and Securities is effective at measuring portfolio concentration of unit portfolios. Further research needs to be done for non-linear portfolios (options) using sensitivity analysis with respect to implied volatility risk factors. Unit portfolios of delta-hedged options are effective at isolating implied volatility risk factors, hence account exceedances attribution to implied volatility can be done through the factor betas with respect to the delta-hedged options.

Several metrics have been proposed in the paper to provide factor directional exceedance contribution. It should be noted that the selection of unit portfolios requires judgement, and the framework is best suited for a limited number of factors. A multi-factor approach with a large number of unit portfolios may double count the expected number of exceedances.

An interesting direction of future research is to connect the expected number of exceedances to aggregate margin model performance using the distribution profiles of clearing member risk profiles. There is some initial analysis of distributions of margin and stress loss risk at CCPs. The analysis by Lewandowska and Mai (2018) builds on the work by Murphy and Nahai-Williamson (2013) with both hypothetical and empirical distributions of CCP account data.

References

Lykourgos Alexiou, Amit Goyal, Alexandros Kostakis, and Leonidas Rompolis. Pricing event risk: evidence from concave implied volatility curves. *Review of Finance*, 29(4):963–1007, 2025. doi: 10.1093/rof/rfaf016. URL https://doi.org/10.1093/rof/rfaf016.

BIS. Framework for supervisory stress testing of central counterparties (ccps), 2018. URL "https://www.bis.org/cpmi/publ/d176.pdf".

Ismael Alexander Boudiaf, Martin Scheicher, and Francesco Vacirca. CCP initial margin models in Europe. Technical report, European Central Bank, 2023. URL "https://www.ecb.europa.eu/pub/pdf/scpops/ecb.op314~afc6d2980c.en.pdf".

Peter Christoffersen. Evaluating Interval Forecasts. *International Economic Review*, 39(4):841–862, 1998.

CPMI. Principles of Financial Market Infrastructures, 2014. URL "https://www.bis.org/cpmi/publ/d101.htm".

Shengwu Du and Travis D. Nesmith. Portfolio Margining Using PCA Latent Factors. Technical Report 2025-016, Board of Governors of the Federal Reserve System (U.S.), 2025. URL https://www.federalreserve.gov/econres/feds/portfolio-margining-using-pca-latent-factors.htm.

Elena-Ivona Dumitrescu, Christophe Hurlin, and Vinson Pham. Backtesting Value-at-Risk: From Dynamic Quantile to Dynamic Binary Tests. *Finance*, 33(2):79–112, 2012.

Robert F. Engle and Simone Manganelli. CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. *Journal of Business & Economic Statistics*, 22(4):367–381, 2004.

ESMA. Opinion: CCP back testing requirements under Article 49 of EMIR. Technical report, European Securities and Markets Authority, 2023. ESMA91-372-2811, 16 Oct 2023.

Pedro Gurrola-Perez. The validation of filtered historical value-at-risk models. *Journal of Risk Model Validation*, 12(2):1–21, 2018. doi: 10.21314/JRMV.2018.185. URL https://www.risk.net/journal-of-risk-model-validation/5455781/the-validation-of-filtered-historical-value-at-risk-models.

Pedro Gurrola-Perez and David Murphy. Filtered historical simulation value-at-risk models and their competitors. Bank of England working papers 525, Bank of England, Mar 2015. URL https://ideas.repec.org/p/boe/boeewp/0525.html.

Christophe Hurlin and Christophe Perignon. Margin Backtesting. 2011. URL https://shs.hal.science/halshs-00746274v1/document.

Philip Kupiec. Techniques for Verifying the Accuracy of Risk Measurement Models. *Journal of Derivatives*, 3(2):73–84, 1995.

LCH. On the margin: Portfolio margining at a CCP. Technical report, London Clearing House, 2023. URL https://www.lseg.com/content/dam/post-trade/en_us/documents/lch/white-papers/on-the-margin.pdf.

Olga Lewandowska and Edgar Mai. The distribution of clearing members risk exposure and how it matters. *Journal of Financial Market Infrastructures*, 7(2):1–24, 2018. URL https://www.risk.net/journal-of-financial-market-infrastructures/6171416/the-distribution-of-clearing-members-risk-exposure-and-how-it-matters.

David Li, Fernando Cerezetti, and Roy Cheruvelil. Correlation breakdowns, spread positions and central counterparty margin models. *Journal of Financial Market Infrastructures*, 11(3):1–22, 2024. doi: 10.21314/JFMI.2024.006.

David Lynch, Iftekhar Hasan, and Akhtar Siddique. *Validation of Risk Management Models for Financial Institutions: Theory and Practice*. Cambridge University Press, Cambridge, 2023.

David Murphy. What can we expect from a good margin model? observations from whole-distribution tests of risk-based initial margin models. LSE Research Online Documents on Economics 118281, London School of Economics and Political Science, LSE Library, Jun 2023. URL https://ideas.repec.org/p/ehl/lserod/118281.html.

David Murphy and Paul Nahai-Williamson. Dear prudence, won't you come out to play? approaches to the analysis of central counterparty default fund adequacy. Technical report, Bank of England, 2013. URL "https://www.bankofengland.co.uk/-/media/boe/files/financial-stability-paper/2014/dear-prudence-wont-you-come-out-to-play-approaches-to-the-analysis-of-ccp-default-fund-adequacy.pdf".

5 Appendix

We provide a brief comparison of existing methods for backtesting margin models in Table 7.

Method	**Type**	**Uses Magnitude?**	**Fixed Portfolio Required?**	**Links Unit → Account?**	**Key Strength / Limitation**
Kupiec (POF)	Unconditional coverage	No (binary)	Yes (strongly)	No	Simple; low power; ignores clustering
Christoffersen (1998)	Conditional coverage	No (binary)	Yes	No	Adds independence; still ignores severity
Basel Traffic Light	Regulatory heuristic	No	Yes	Weak (portfolio only)	Easy to implement; not statistical
Time Between Failures (TUFF / Haas)	Duration test	No	Yes	No	Captures clustering; still binary
Dynamic Quantile (DQ) (Engle–Manganelli)	Regression-based	No (binary)	Not strictly	Limited	Detects conditional misspecification
Pearson Q / Multi-quantile tests	Distributional	Partial	Yes	No	Improves power vs Kupiec; still weak for tails
Berkowitz Test	Density / PIT test	Yes (full distribution)	Yes	No	Uses full distribution; sensitive to misspecification
Lopez Loss Function	Loss-based	Yes	Not required	Weak	Incorporates magnitude; ad hoc
APL (Quantile Loss / Pinball)	Proper scoring rule	Yes (piecewise linear)	Not required	Yes (extendable)	Captures severity + frequency; consistent estimator
Fissler–Ziegel (VaR+ES scoring)	Joint scoring	Yes	Not required	Yes	Regulatory direction (ES); more complex
Bootstrap / Simulation-based tests	Resampling	Yes	Not required	Possible	Good finite sample properties; computational
Factor-based Backtesting (research)	Structural / regression	Yes	Not required	**Yes (direct)**	Links failures to factors; not standardized

Table 7: Comparison of VaR Backtesting Methods and Their Applicability

Symbol	**Name**	**Definition**
$\beta_{a,k}$	Factor beta	$\frac{Cov(\Delta V, \Delta f_k)}{Var(\Delta f_k)}$ - portfolio sensitivity to factor k
$\hat{q}$	Unit portfolio exceedance rate	Fraction of days factor k breaches its VaR
$\hat{p}$	Account exceedance rate	Fraction of days the account breaches IM
$\hat{p}_{a\|k}$	Conditional exceedance rate	P(account breach| factor k breach)
$C_{a,k}$	Concentration Ratio	how concentrated breaches are on factor k
$C^d_{a,k}$	Directional concentration	concentration ratio restricted to adverse factor moves
$A_{a,k}$	Asymmetry ratio	Adverse vs favorable conditional exceedance rates
$\phi_{a,k}$	Magnitude attribution	fraction of exceedance loss on factor k

Table 8: Glossary of Notation

Condition	**Targeted Add-on Appropriate**	**Model Recalibration Required**
Backtesting failures localized to specific spreads or factors	Yes apply a factor-level or portfolio-level add-on	No
Failures concentrated in specific accounts (concentration risk)	Yes account-level add-on based on exposure weights	No
Systematic backtesting failures across most portfolios	No	Yes recalibrate model
Persistent underestimation of volatility across market states	No	Yes revise model dynamics (e.g., FHS decay)
Correlation breakdown affecting broad product class	Possibly (temporary)	Yes recalibrate correlation structure
Procyclicality concerns (margin cliff effects)	No	Yes apply system-wide APC tools
Tail Risk in option portfolio (Convexity)	Yes	Severity analysis of shortfall beyond margin
Model misspecification (distributional / structural)	No	Yes model redevelopment required

Table 9: Recommendations: Targeted Add-ons vs Model Recalibration