

Real Contracts under Common Trading Scenarios

A Comparison Framework with a Source-Linked Pilot

Francisco Matilla Serrano

Independent Researcher, Varna, Bulgaria

Corresponding author: Francisco Matilla Serrano (franchimatilla@gmail.com)

ORCID: <https://orcid.org/0009-0004-9491-6374>

18 September 2026

Abstract

Ranking trading evaluations by price or a target-to-drawdown ratio assumes that omitted contract terms do not change the ordering. This paper defines a comparison that instead holds trading scenarios fixed while replaying versioned contracts. It separates evaluation passing, funded mechanical eligibility, actual payment and participant net cash flow, and specifies how fees and conditional activation enter each path. The V1.8 Atlas contains 177 product-route-size configurations across 22 officially coded providers. Of these, 155 evaluation configurations and 103 funded configurations contain the inputs required by the current event engine. The application replays EOD, intraday and static loss floors, soft daily loss limits, consistency thresholds, minimum trading days, qualifying days and coded first-request hurdles under three common synthetic P&L scenarios. It produces 774 stage-scenario estimates, with paired paths, standard errors and Wilson intervals retained in the accompanying workbook. The paper derives an exact pairwise value decomposition and an approval-sensitivity boundary, then defines validation and multiplicity rules for named comparisons. It does not estimate observed provider pass rates, actual payment reliability or a universal best provider. The numerical results are counterfactual outputs conditional on documented rules and declared scenarios; incomplete rule fields remain excluded rather than imputed.

Keywords: contract comparison; common random numbers; payout risk; Monte Carlo; identification

JEL Classification: C15; D81; D86; G23

1. Research question and scope

Which real contract is more favorable under the same trading scenario, and how stable is that comparison across plausible scenarios? This is not a question about which provider has the highest observed success rate. Providers may attract different traders and use different counting conventions. Nor is it the synthetic intervention in *The Rules Are the Model*, which changes selected rules while holding an artificial benchmark fixed.

The present question compares complete observed packages. A product may be less expensive but have a tighter floor, a lower request cap or a different funded phase. The relevant output is conditional on a declared trader law, strategy and horizon. There need not be one winner across all conditions.

The contractual atlas supplies the sampling frame and source records. The V1.8 snapshot covers 177 product-route-size configurations across 22 officially coded providers, with one discovered provider unresolved. This is a broad futures-only corpus, not a claim to have enumerated every contract sold worldwide. The application retains provider-product-route-size-version identity and does not combine an old evaluation target with a current funded cap. Unknown rules are a reason to withhold a complete score, not an invitation to fill them with favorable defaults.

Lim's *The Price of a Funded Account* (2026b), Lim's *Phantom Generosity* (2026a), and Tomàs Fernández (2026) already establish single-contract and cross-firm valuation benchmarks. The specific contribution here is to replay the versioned futures-only Atlas on paired common paths, report the evaluation and funded denominators separately, and withhold estimates when an event cannot be reconstructed from the coded rules. Within the program, it is the step that converts the Atlas from a documentary input into a reproducible product-level counterfactual comparison.

2. Define the outcome before simulating

For contract c and simulated path i , let E_{ic} indicate a pass. Let Y_{ic} indicate satisfaction of the modeled funded conditions following that pass, so Y_{ic} cannot exceed E_{ic} . Let R_{ic} indicate that an eligible claim is actually paid. A pathwise first-receipt cash flow is

$$C_{ic} = -F_c - A_c E_{ic} - H_{ic} + Y_{ic} R_{ic} W_{ic}.$$

F is the initial fee, A the activation charge, H subsequent charges included in the selected horizon, and W the participant receipt after the contract's share and caps. If W is already net of the share, it must not be multiplied by that share a second time. A buffer affects feasible W or Y ; it is not automatically a cash fee.

The primary estimand is the expectation of C conditional on the declared scenario. Secondary outcomes include pass probability, eligibility probability from funded entry, receipt probability, time to event and probability of strict net profit. A study must not call all of these 'success'. It must also state whether an unfinished account has residual value or simply lies beyond the reporting horizon.

3. Pairwise value decomposition

In a simplified one-attempt comparison, write G_c as expected participant receipt from a fresh funded account and p_c as evaluation-pass probability. With initial fee F_c and activation A_c ,

$$V_c = p_c(G_c - A_c) - F_c.$$

For contracts c and d , an exact decomposition is

$$V_c - V_d = (p_c - p_d)(G_d - A_d) + p_c[(G_c - A_c) - (G_d - A_d)] - (F_c - F_d).$$

The first term values the access-probability difference at contract d 's continuation value. The second captures the different funded opportunity and activation price. The third is the entry-price

difference. Reversing the reference contract produces another valid decomposition, so the component attribution is reference-dependent even though the total is not.

This identity explains why headline price cannot establish a value ranking. It also prevents calling a high pass probability beneficial without considering whether entering the funded stage has positive net continuation value. The application must estimate G rather than substitute the advertised maximum payout.

4. Payment is not identified by the rulebook

Suppose the engine establishes mechanical eligibility q_c and conditional eligible participant receipt w_c from funded entry. Actual approval and transfer reliability may still be unknown. Under the explicitly simplified scalar payment probability r_c , expected receipt is $q_c r_c w_c$. If a nonconstant payment process depends on account state or amount, the appropriate quantity is an expectation of the product instead.

The scalar formulation supplies a useful sensitivity boundary. For $p_c q_c w_c > 0$, first-attempt value is positive only when

$$r_c > \frac{F_c + p_c A_c}{p_c q_c w_c}.$$

If the right-hand side exceeds one, even certain payment cannot make that calibration positive. If it lies below one, the formula states the reliability required; it does not establish that a real provider meets it. A paper should present an unidentified parameter as a sensitivity dimension rather than assign a convenient empirical-looking number.

5. What the V1.8 application demonstrates

Some matched 50K Atlas routes have $U=3,000$ and $D=2,000$ during evaluation. The ratio U/D is therefore 1.5 for that subset. A static-barrier benchmark ties those products for any common μ , σ and horizon, whereas the V1.8 event engine can separate them when their trailing mode, daily loss limit, consistency rule or minimum-day condition differs. A tied headline metric cannot recover distinctions it does not encode.

The funded comparison is not tied even at the descriptive level. A retained buffer can require more account profit before a given request; a 50% withdrawal fraction can require twice the requested gross amount. Some plans add qualifying days; some change from EOD to intraday trailing. These documented differences establish that the mapping from headline geometry to the complete contract is not one-to-one.

The model establishes the size and ordering of scenario-conditional mechanical events under the implemented rules. It still does not establish complete participant value because prices, activation, renewal, review discretion, payment reliability and behavioral adaptation are not all identified for every product. It would therefore be misleading to describe a probability ordering as a validated provider ranking. The application reports transparent product-level estimates and prespecified contrasts, not a league table.

6. Full-engine implementation protocol

The first-passage literature, including Magdon-Ismail et al. (2004) on drawdown, informs the stochastic framework but does not encode these provider-specific conditions. The engine should take a versioned rule object, a P&L scenario and an event-order specification. State must include current balance, high-water mark, loss floor, best daily profit, completed days, qualifying-day count, stage, request history and any tier determining exposure or daily limits. A rule that changes after payout requires a state transition, not a permanently fixed scalar.

Table 1. Required model-to-contract checks

Test	Example of what must be checked
Boundary touch	Whether equality causes failure
EOD ratchet	Intraday gain does not raise an EOD-only floor
Intraday ratchet	Open profit can tighten the loss floor
Daily stop	Pause differs from permanent termination
Consistency	Denominator, equality and reset time are correct
Withdrawal	Buffer and cap are applied in the stated order
Fees	Activation occurs after every activated pass
Version	Purchase cohort selects the intended rule object

Deterministic hand paths should test every rule before large simulation. For example, a path that earns strongly intraday and gives back profits by the close distinguishes EOD from intraday ratchets. A request made just below and just above the minimum distinguishes eligibility from rounding. A second request after a losing cycle tests whether accumulated historical profit is sufficient or new cycle profit is required.

6.1 Estimable universe and matched comparisons

Every catalogue row is eligible for descriptive comparison, but only rows with a verified target, loss budget, drawdown timing and required qualification fields enter the evaluation simulation. Funded-stage estimates additionally require a defined mechanical eligibility hurdle. This creates three denominators: all coded configurations, evaluation-modelable configurations and funded-modelable configurations. Reporting only the last denominator would hide where the rulebook is incomplete.

Comparisons are matched on nominal size only when size is economically relevant to the stated process. The main synthetic scenarios scale drift and volatility by the initial loss budget, which isolates rule geometry more directly than scaling by the nominal account label. A complementary dollar-P&L scenario holds the trader process fixed across 50K products and shows how nominally similar accounts respond to different actual loss budgets.

The primary pairwise estimand for evaluation is

$$\Delta_{cd}^E(s) = P_s(E_c = 1) - P_s(E_d = 1),$$

for contracts (c) and (d) under scenario (s). The funded counterpart is defined for first-request mechanical eligibility conditional on funded entry. Neither estimand is an observed provider outcome rate. A complete participant-value comparison would also add verified prices, activation, payout share, payment reliability and re-entry policy.

6.2 Event engine and deterministic tests

The rule engine updates, in order, the session opening state, intraday P&L, high-water mark, active floor, daily stop, permanent breach, session-close statistics and qualification conditions. This order is part of the contract specification. EOD and intraday trailing cannot be represented by the same floor update, and a soft daily pause cannot be represented as permanent failure.

Before simulation, every rule module is exercised on hand-built paths. At minimum, tests cover equality at target and loss boundaries, intraday profit followed by an EOD giveback, a daily-stop hit without total failure, a consistency failure that later becomes compliant, minimum-day completion after an early target touch, and a funded request exactly at the buffer threshold. A product enters the named comparison only after its applicable tests pass.

7. Common paths and uncertainty

For a fixed trading policy, common simulated shocks should be replayed across contracts. Pairwise differences are then measured on the same scenario realizations. If $d_i = C_{ic} - C_{id}$, the Monte Carlo standard error uses the sample standard deviation of d_i divided by the square root of the number of independent simulation units. Treating each contract's estimate as independent discards covariance that the design intentionally creates.

Common shocks do not mean identical realized P&L after a rule stops trading or limits exposure. The engine must branch state and execution when the contract requires it, while keeping the underlying opportunity sequence common. This makes the comparison a clearly defined counterfactual experiment, not a claim that real traders would behave identically under every contract.

Policy adaptation is a separate scenario. One comparison can hold the trading policy fixed; another can allow each contract its best policy within the same declared library. Selecting that policy on training paths and evaluating on fresh holdout paths separates optimization noise from reported performance. Neither design should be mislabeled as the other.

7.1 Uncertainty and multiplicity

Common random numbers make pairwise contrasts more precise because each contract sees the same underlying opportunity path until its rules alter trading state. Standard errors are calculated from paired path-level differences. Wilson intervals remain useful for a single binomial probability, but subtracting two independent Wilson intervals is not the correct uncertainty calculation for a common-path contrast.

A market-wide table contains many possible pairwise comparisons. The analysis therefore avoids declaring a winner from all $(K(K-1)/2)$ contrasts. The primary comparisons are prespecified rule contrasts: EOD versus intraday trailing within the same family and size; with versus without a soft DLL; and products with the same target and loss budget but different

minimum-day or consistency rules. Holm adjustment is applied within each declared family when formal decisions are reported.

8. Robust rankings rather than one league table

A contract that has higher value under every declared scenario dominates within that scenario set. Otherwise, report the regions in which the ordering reverses. The scenario set should vary drift, volatility, timing, dependence and relevant fee durations. Results should not be selected after seeing which provider wins.

Unknown terms can be bounded where genuine contractual limits are known. If a field can take a broad range with no defensible endpoints, a ranking may simply be unidentified. An honest incomplete comparison is more informative than an apparently precise score determined by undisclosed assumptions.

8.1 Scenario dominance and ranking stability

A contract dominates another within the declared scenario set only if its participant-value estimate is no lower in every scenario and strictly higher in at least one. If the ordering reverses with drift, volatility, horizon or process scaling, the output is a ranking map rather than one league table. The number and location of reversals are themselves results because they show which assumptions drive the comparison.

Unknown rules are handled by bounds where defensible endpoints exist. If a funded hurdle is known to lie between a documented buffer and buffer plus a capped request, both endpoints can be run. If no credible range is available, the comparison is unidentified and remains excluded. This rule prevents missing evidence from being converted into an apparently precise provider score.

9. V1.8 application and interpretation

The expanded Atlas supplies a much larger official-source product universe than the original eight-provider pilot. The accompanying workbook calculates scenario-conditional probabilities for every configuration whose required geometry is closed. The manuscript summarizes rule contrasts and coverage; the workbook retains the product-level estimates, standard errors, Wilson bounds, simulation counts, horizon, monitoring grid and seed.

These estimates improve the comparison from a headline target-to-drawdown ratio to a versioned rule engine. They do not complete the payment side of participant value. Review discretion, undocumented execution effects, taxes, slippage and actual behavioral adaptation remain outside the model. Named-provider probability results must therefore be read as counterfactual outputs under the declared synthetic P&L scenarios, not predictions of what that provider's customers achieve.

9.2 Completed product-level comparison

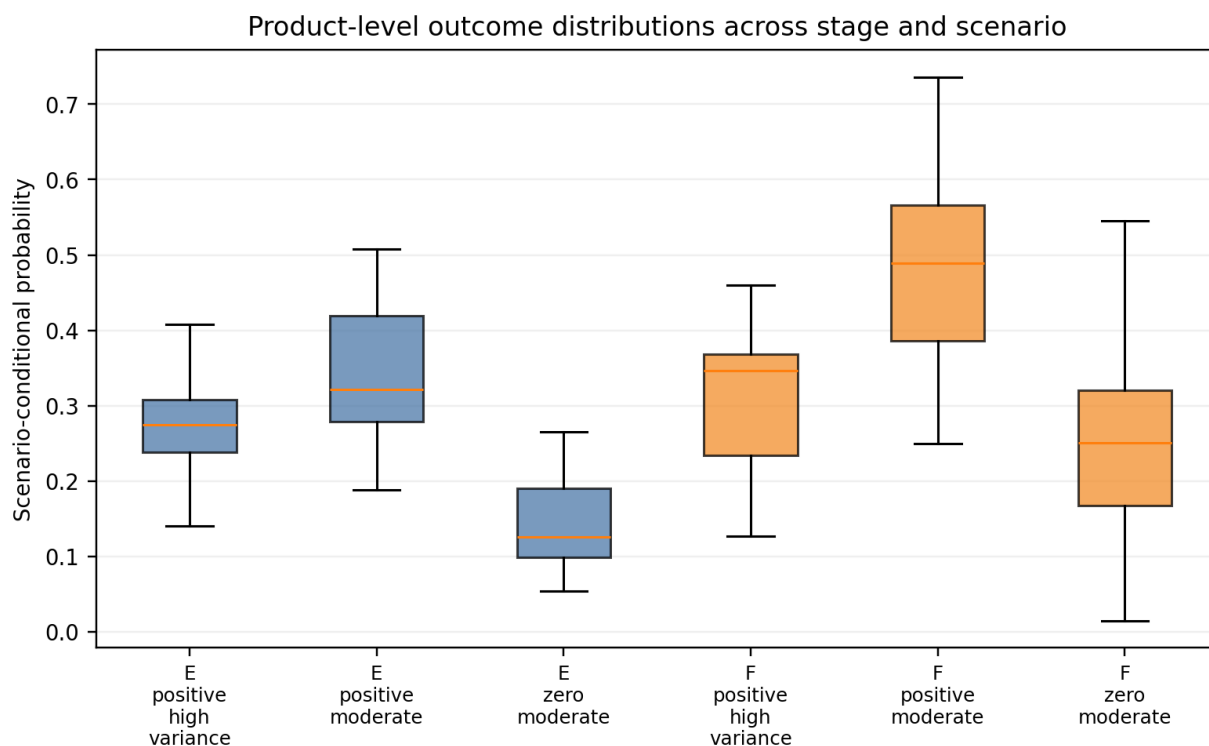
The V2.0 application retains all 774 modelable product-stage-scenario rows rather than reducing the Atlas to one 50K illustration. The 465 evaluation rows and 309 funded-stage rows identify provider, family, route, nominal size, scenario, simulation count, event count, probability,

Monte Carlo standard error, Wilson interval, horizon, monitoring grid and seed. Provider summaries are descriptive views of this product-level file, not replacements for it.

Table 4. Distribution of modelled probabilities across products

Stage	Scenario	Products	Median	P10	P90
Evaluation	positive high variance	155	27.50%	22.99%	36.67%
Evaluation	positive moderate	155	32.18%	27.36%	42.90%
Evaluation	zero moderate	155	12.53%	9.70%	19.70%
Funded	positive high variance	103	34.68%	14.23%	41.27%
Funded	positive moderate	103	48.90%	26.70%	68.38%
Funded	zero moderate	103	25.09%	10.05%	46.42%

The spread within each scenario is economically material, but it is not an observed league table. Products differ in target, loss budget, trailing convention, daily limits, minimum days and payout hurdles. The workbook therefore provides matched extracts by stage, scenario, nominal size and rule family. A named provider's estimate remains conditional on the declared synthetic P&L process.



Product-level outcome distributions by stage and scenario.

Figure 1. Product-level outcome distributions by stage and scenario. Boxes summarize the declared synthetic model outputs; they are not historical provider success rates.

Table 5. Positive-moderate scenario medians by provider and stage

Provider	Evaluation median (products)	Funded median (products)
Apex Trader Funding	34.86% (8)	47.60% (2)
AquaFunded Futures	36.13% (4)	Not modelled
Blue Guardian Futures	39.19% (8)	43.44% (4)
Blueberry Futures	34.98% (8)	Not modelled
Day Trader Funding	29.34% (3)	Not modelled
DayTraders	33.56% (4)	25.25% (4)
E8 Futures	35.75% (4)	36.47% (4)
Earn2Trade	26.09% (1)	Not modelled
Funded Futures Family	28.61% (12)	57.34% (4)
FundedNext Futures	36.62% (12)	Not modelled
FuturesElite	32.97% (4)	63.43% (4)
Goat Funded Futures	28.58% (9)	26.61% (13)
Lucid Trading	33.34% (28)	48.57% (24)
My Funded Futures	28.58% (11)	49.95% (5)
Take Profit Trader	41.94% (1)	49.95% (5)
The Trading Pit Futures	27.75% (3)	40.90% (3)
The5ers Futures	38.17% (1)	58.82% (1)
Top One Futures	33.30% (8)	40.31% (8)
Topstep	28.55% (3)	69.52% (3)
TradeDay	27.20% (6)	68.00% (6)
Tradeify	35.28% (14)	56.32% (10)
Traders Launch	27.76% (3)	58.82% (3)

The parenthetical number is the count of product configurations contributing to the median. A missing funded result means that the required funded-stage event geometry was not closed for that provider in the current model, not that the probability is zero.

9.3 Ranking stability and sensitivity

Provider-level ordering is recalculated separately by scenario and stage. The rank-stability table reports the common provider set and the mean absolute change in position between adjacent scenarios. Reversals are retained rather than hidden by one aggregate score. This is important because a rule set that is favorable to a positive-drift, moderate-volatility process need not remain favorable under zero drift or higher volatility.

The main comparison scales trading drift and volatility to the effective loss budget. The 50K extract provides a complementary dollar comparison. Neither scaling convention is uniquely correct, so agreement across them is stronger evidence than a result that appears under only one convention. Fee, activation, payout share and payment reliability remain separate inputs; missing commercial terms are not imputed.

10. Limitations and conclusion

The V1.8 application contains an executable rule engine for the coded evaluation and funded-stage mechanics, but it is not a model of every contractual or operational contingency. Products with unresolved core inputs are excluded. Review discretion, actual approval and transfer, slippage, taxes, strategy adaptation and the distribution of real participant behavior remain outside the current estimates. The application can therefore report named products only as source-linked counterfactual cases under the stated scenarios, not as observed performance or a recommendation.

Equal headline geometry does not imply equal contractual opportunity, and passing is only one component of participant value. The contribution reported here is the 774-row, source-linked application under three declared scenarios, together with paired-path uncertainty and explicit model-coverage exclusions. Matilla Serrano (2026a) identifies the contracts; this paper measures the conditional consequences of the modelable rule packages while leaving actual customer behavior and payment reliability unidentified.

References

- Lim, B. C. (2026a). *Phantom Generosity: Contract Design and Value in the Market for Proprietary-Trading Evaluations*. SSRN preprint, posted 5 August 2026. <https://doi.org/10.2139/ssrn.7184138>
- Lim, B. C. (2026b). *The Price of a Funded Account: An Actuarial Analysis of Proprietary Trading Firm Challenges*. SSRN preprint, posted 5 August 2026. <https://doi.org/10.2139/ssrn.7178078>
- Matilla Serrano, F. (2026a). *A Versioned Contract Atlas of Retail Futures Evaluations: Product Families, Stage Transitions, and Scenario-Conditional Probabilities*. SSRN preprint, posted 14 September 2026. <https://doi.org/10.2139/ssrn.7429100>.
- Magdon-Ismael, M., Atiya, A. F., Pratap, A., and Abu-Mostafa, Y. S. (2004). On the maximum drawdown of a Brownian motion. *Journal of Applied Probability*, 41(1), 147-161. <https://doi.org/10.1239/jap/1077134674>
- Tomàs Fernández, M. (2026). *Valuing Proprietary Trading Firm Evaluation Contracts: Closed Form, Cross-Firm Dispersion and the Source of the Margin*. SSRN preprint, posted 12 August 2026. <https://doi.org/10.2139/ssrn.7260819>

The documentary examples use the official sources listed in the atlas of Matilla Serrano (2026a). No practitioner simulator outputs are used as observations.

Declarations

Status. Source-linked counterfactual comparison under declared synthetic scenarios, not an observed provider-outcome study. No real pass rate, payment rate or universal provider ranking is claimed.

Conflicts of interest. No financial or commercial ties exist with any provider included in the comparison.

Supplementary materials. The source-linked contract data, comparison workbook, scenario definitions, analysis code, and reported numerical outputs are available on Zenodo: <https://doi.org/10.5281/zenodo.22770522>.