

Risk Is a Choice

Optimal Trading under Evaluation Constraints

Francisco Matilla Serrano

Independent Researcher, Varna, Bulgaria

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

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

19 September 2026

Abstract

Provider rules change evaluation outcomes, but participants also respond to those rules. This paper studies risk choice inside the five fictitious 50K contract systems introduced earlier in this research program: Ideal Basic (I0), Ideal Realistic (I1), Static Discipline (A), EOD Trail (B), and Intraday Guardrail (C). Every system retains a \$3,000 evaluation target, \$2,000 headline drawdown, and 20-session horizon. Contracts A-C then add minimum days, consistency, daily-loss, trailing-drawdown, buffer, and payout-cap rules.

The analysis has two layers. For the static contract, a daily-grid dynamic program supplies an upper numerical benchmark within a finite action set. For the path-dependent contracts, a global optimum would require a much larger state space. The paper therefore compares nine declared policies, selects the best policy on 12,000 training paths, and reports it on 30,000 independent holdout paths. “Best response” for A-C means best in that finite policy library, not globally optimal trading.

For a positive-drift, lower-variance trader, fixed baseline risk passes 45.06% under I0 and I1, 43.90% under A, 41.55% under B, and 20.26% under C. The training-selected pass policy raises the figures to 60.64%, 60.64%, 50.60%, 41.55%, and 20.26%, respectively. The static dynamic-programming policy transfers well to I0 and I1, fixed high risk is selected for A, and baseline risk remains best in the tested library for B and C. Rules do not merely subtract a constant probability; they change the participant’s response.

The paper also reconciles an apparent difference across the research line. Matilla Serrano (2026c) checked both target and loss on a ten-point intraday grid. The original dynamic program in this paper checked both at daily closes. The rule engine in Matilla Serrano (2026d) checks losses intraday but determines qualification at the close. These are different stopping rules, so their numerical pass probabilities need not match. Monitoring is part of the contract model.

The extended analysis adds an 80-cell drift-policy-contract surface and 20 fresh holdout replications for each contract. These results test computational robustness under the declared synthetic process; they do not introduce market observations.

Recent preprints already analyze contract value, first-passage probabilities, and risk sizing in funded-account evaluations (Lim, 2026a, 2026b; Tomàs Fernández, 2026). The contribution here is the controlled policy comparison after risk becomes a participant choice. Within the program, that comparison is necessary because a contract rule can change both the feasible path and the response selected by the trader.

Keywords: trading evaluation; stochastic control; risk choice; best response; trailing drawdown; consistency rule; finite horizon

JEL Classification: C15; C61; D81; G11; G23

1. Introduction

The Rules Are the Model held P&L paths fixed and changed the contract (Matilla Serrano, 2026d). The natural complement is to let the participant change exposure after seeing the contract. A daily-loss limit can make high risk unattractive. A deadline can make low risk ineffective. A consistency rule can punish one concentrated gain. A trailing boundary can turn an earlier profit into a tighter future loss floor. The participant's risk decision is therefore contract-dependent.

The advertised nominal account is 50K, but the economic state is not a freely withdrawable \$50,000 portfolio. It is cumulative P&L relative to a target, loss boundary, rule state, and deadline. The evaluation fee is sunk after purchase; activation is paid after a pass; payout depends on surviving a second stage and satisfying payout rules.

A static-barrier model is useful for explaining the dynamic-programming logic, but by itself it is not a full counterpart to Matilla Serrano (2026d). The analysis therefore applies the same participant-choice exercise to I0, I1, A, B, and C. It asks:

1. Which risk policy performs best under each synthetic contract?
2. Does a policy that works under the ideal contract remain attractive after adding minimum-day, consistency, daily-loss, and trailing rules?
3. How do evaluation-pass and first-payout probabilities move together?
4. Does the policy selected for pass probability also maximize expected participant cash flow in the tested library?
5. Why do pass probabilities differ across the selection, risk-control and robustness implementations even when drift, volatility, target, and drawdown appear identical?

The answer to the first four questions is empirical within the simulation. The answer to the fifth is definitional: a barrier event depends on when the barrier is observed and when qualification becomes final.

2. Place in the Research Program

Market Edge versus Contract Edge separated expected trading P&L from the cash flows created by fees, activation, and payouts (Matilla Serrano, 2026b). *Skill or Variance?* made pass probability endogenous to drift, volatility, and a finite horizon (Matilla Serrano, 2026c). *The Rules Are the Model* replayed common paths through alternative rule packages (Matilla Serrano, 2026d).

This fourth paper keeps those contract packages and changes the participant policy. It is therefore a strategic response exercise, but not a game between provider and participant. Provider rules are fixed. The participant chooses only a daily exposure multiplier.

Goal-reaching under a deadline is an established stochastic-control problem (Browne, 1999). Drawdown control is also an established portfolio problem (Cvitanic and Karatzas, 1995; Chekhlov, Uryasev, and Zabaranin, 2005; Nystrup et al., 2019). Nonlinear rewards can induce state-dependent risk shifting (Carpenter, 2000; Basak, Pavlova, and Shapiro, 2007). The question here is how those ideas operate inside the five two-stage evaluation systems used

across this research line. A fixed policy, a small library of adaptive rules, and a static dynamic program are compared on common training and holdout paths.

The direct trading-evaluation literature provides the valuation baseline. Lim (2026a, 2026b) studies market contracts and a stylized actuarial challenge, and Tomàs Fernández (2026) derives closed-form values and break-even conditions. This paper does not claim to introduce valuation or risk sizing. It isolates a different estimand: the best tested response within a declared policy library, selected on training paths and evaluated on fresh holdout paths under each synthetic rule system.

3. Contract Systems

All systems use the same nominal label and headline barriers. I0 and I1 have identical mechanical rules; they differ only in fee, activation, and payout share. Contracts A-C use the anonymized synthetic bundles defined in Matilla Serrano (2026d).

Table 1. Synthetic contract systems

ID	Fee	Act. fee	Share	Floor	Daily loss	Min. days	Cons.	Buffer	Cap
I0	\$40	\$0	100%	Static	n/a	1	n/a	\$0	\$3,000
I1	\$90	\$100	90%	Static	n/a	1	n/a	\$0	\$3,000
A	\$90	\$100	90%	Static	n/a	5	50%	\$500	\$2,000
B	\$90	\$100	90%	EOD trailing	\$1,500	5	50%	\$1,000	\$2,000
C	\$90	\$100	90%	Intraday trailing	\$1,000	7	30%	\$1,000	\$1,500

Table abbreviations: Act. fee = activation fee; Cons. = consistency threshold; n/a = not applicable. All monetary amounts are USD.

Every contract has a \$3,000 target and \$2,000 headline drawdown, so the Target-to-Drawdown Ratio is 1.5 throughout. Differences in results arise from path rules, timing, and transfers, not from the ratio.

The funded stage applies the corresponding minimum days, consistency, trailing, daily-loss, buffer, cap, and 90% share. Evaluation and funded-stage path panels are independent. Activation is charged after every evaluation pass, independently of whether the funded stage later reaches payout.

4. Controlled P&L and Policy Library

4.1 Baseline trader

The baseline is the positive-edge, lower-variance type used in Matilla Serrano (2026c, 2026d):

$$\mu = \$100 \text{ per session}, \sigma = \$600 / \sqrt{\text{session}}.$$

Ten intraday shocks make up each session. The participant selects one risk multiplier a_d at the start of day d , and that multiplier applies to all ten increments:

$$\Delta X_{d,k} = a_d \left(\frac{\mu}{10} + \frac{\sigma}{\sqrt{10}} Z_{d,k} \right), Z_{d,k} \sim N(0, 1).$$

Risk can change daily but not inside the session. This preserves a manageable and interpretable control while allowing intraday loss monitoring.

4.2 Nine policies

The policy library contains:

1. fixed low risk, $a = 0.50$;
2. fixed baseline risk, $a = 1.00$;
3. fixed high risk, $a = 1.50$;
4. a barrier-aware rule;
5. a deadline-aware rule;
6. a cautious rule-aware policy;
7. a balanced rule-aware policy;
8. an aggressive rule-aware policy; and
9. the static-contract dynamic-programming policy transferred to the rule engine.

The rule-aware policies react to distance from the current floor, distance from target, sessions remaining, minimum-day status, and consistency pressure. Their thresholds are fixed before training. Training chooses among policies; it does not tune a large continuous parameter set.

4.3 Static dynamic-programming benchmark

For a static contract observed once per session, define $V_d(x)$ as the maximum probability of hitting $U = \$3,000$ before the deadline or the lower boundary $-D = -\$2,000$. The action set is

$$A = \{0.35, 0.50, 0.75, 1.00, 1.25, 1.50, 1.75, 2.00\}.$$

Backward induction gives

$$V_d(x) = \max_{a \in A} E[V_{d+1}(x + a\mu + a\sigma Z)].$$

The expectation uses 31-node Gauss-Hermite quadrature on a \$50 state grid. This is a numerical upper benchmark only within the static daily model.

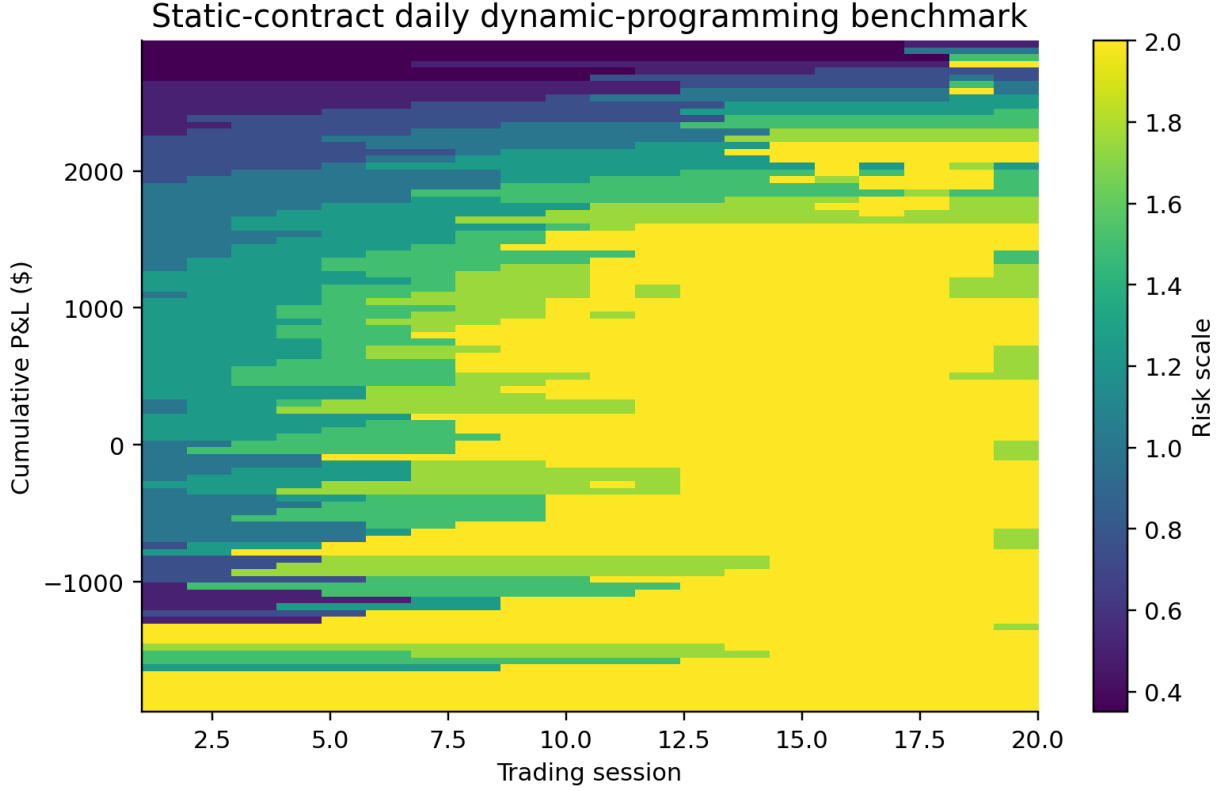


Figure 1. Static-contract daily dynamic-programming risk map.

Trailing drawdown, daily loss, minimum days, and consistency add state variables: current high-water mark, session-start equity, best positive day, days completed, and stage. Rather than hide that dimensional increase, this paper limits its A-C claim to the best tested policy.

4.4 State, action, and feasibility map

An adaptive policy is well defined only if its information set and feasible actions are explicit. At the start of session (d), the state used in this paper can be written as

$$S_d = (X_d, H_d, L_d, B_d, m_d, c_d, \tau_d),$$

where (X_d) is cumulative P&L, (H_d) is the relevant high-water mark, (L_d) is the active loss floor, (B_d) is the largest positive day used by the consistency rule, (m_d) counts completed or qualifying days, (c_d) identifies the contract and stage, and (τ_d) is the remaining horizon. The action is the exposure multiplier (a_d) selected before the session shocks are observed. A policy may use (S_d), but not future returns, the closing P&L of the current session, or the holdout realization.

This state list also explains why a rule-aware problem is larger than the static dynamic program. Under a static floor and no timing restrictions, current P&L and time are sufficient for the simplified probability objective. A trailing floor requires the high-water mark; a daily stop requires the session opening balance; consistency requires the best-day statistic and cumulative profit; minimum-day and payout rules require counters and stage history. Omitting any variable that changes future feasible outcomes makes the numerical policy non-Markov even if the underlying shocks are Markov.

The nine-policy library is deliberately small. It contains three constant exposures, two simple distance or deadline responses, three rules that react to contractual state, and one transferred static dynamic-programming rule. The design tests whether economically interpretable adaptations matter before introducing a high-dimensional optimizer. It should not be read as an exhaustive search or a claim of global optimality. Contract limits, integer futures contracts, commissions and signal capacity would further restrict the action set in an implementation.

Table 2. Policy information and interpretation

Policy family	Uses contract state?	Main variation	Interpretation
Fixed exposure	No	Constant multiplier	Transparent risk-scale benchmark
Barrier aware	Distance to target and floor	Reduces risk near loss boundary	Local survival response
Deadline aware	Target distance and sessions left	Raises exposure when time is scarce	Goal-reaching response
Rule aware	Floor, days and consistency state	Changes exposure around active constraints	Contract-specific heuristic
Static-DP transfer	Static P&L and time state	Grid-selected multiplier	Upper benchmark only for its restricted model

5. Training and Holdout Design

All nine policies are run on 12,000 common antithetic training paths for each contract and stage. Two policies are identified per contract:

- the policy with the highest evaluation-pass probability; and
- the policy with the highest participant expected cash flow.

The selected labels are then frozen and evaluated on 30,000 new antithetic holdout paths. Using a holdout does not make the policy empirically valid, but it prevents the paper from reporting the largest value from the same Monte Carlo noise used for selection.

Participant expected cash flow is

$$EV = p E[\text{participant payout} \mid \text{funded start}] - F - p A.$$

The expected funded-stage payout already contains funded failure, buffer, cap, and share. The evaluation pass probability p determines whether activation is paid and whether the funded-stage lottery is entered.

5.1 Prespecified sensitivity surface

The central result is conditional on the drift-volatility pair and the contract system. A reproducible extension should therefore evaluate the frozen policies over a surface rather than at one point. The prespecified grid is $\mu \in \{0, 50, 100, 150\}$ dollars per session and $\sigma \in \{300, 600, 900, 1,200\}$ dollars per square-root session, crossed with horizons of 10, 20 and 40 sessions. Each cell must use common paths across policies, a training sample for policy selection and a separate holdout sample for reporting.

The surface answers three distinct questions. First, it shows whether a policy advantage is local to the baseline calibration. Second, it reveals whether a rule changes the response to drift or to

volatility. Third, it distinguishes a high pass probability from high participant expected cash flow after activation, payout caps and failure are included. A policy should be described as robust only over the declared region in which it retains its advantage; an isolated winning cell is not evidence of a generally superior trading rule.

No additional numerical surface is inserted into this revision because the existing holdout results were generated for the baseline calibration. The grid above is a locked extension design, not an unreported result. This boundary avoids presenting exploratory numbers as if they had been selected in advance.

6. Results

6.1 Contract-dependent pass response

Table 3. Holdout pass probability and training-selected response

Contract	Fixed baseline	Selected policy	Selected pass	First-payout probability under selected policy
I0	45.06%	Static-DP transfer	60.64%	36.77%
I1	45.06%	Static-DP transfer	60.64%	36.77%
A	43.90%	Fixed high	50.60%	25.70%
B	41.55%	Fixed baseline	41.55%	17.22%
C	20.26%	Fixed baseline	20.26%	4.14%

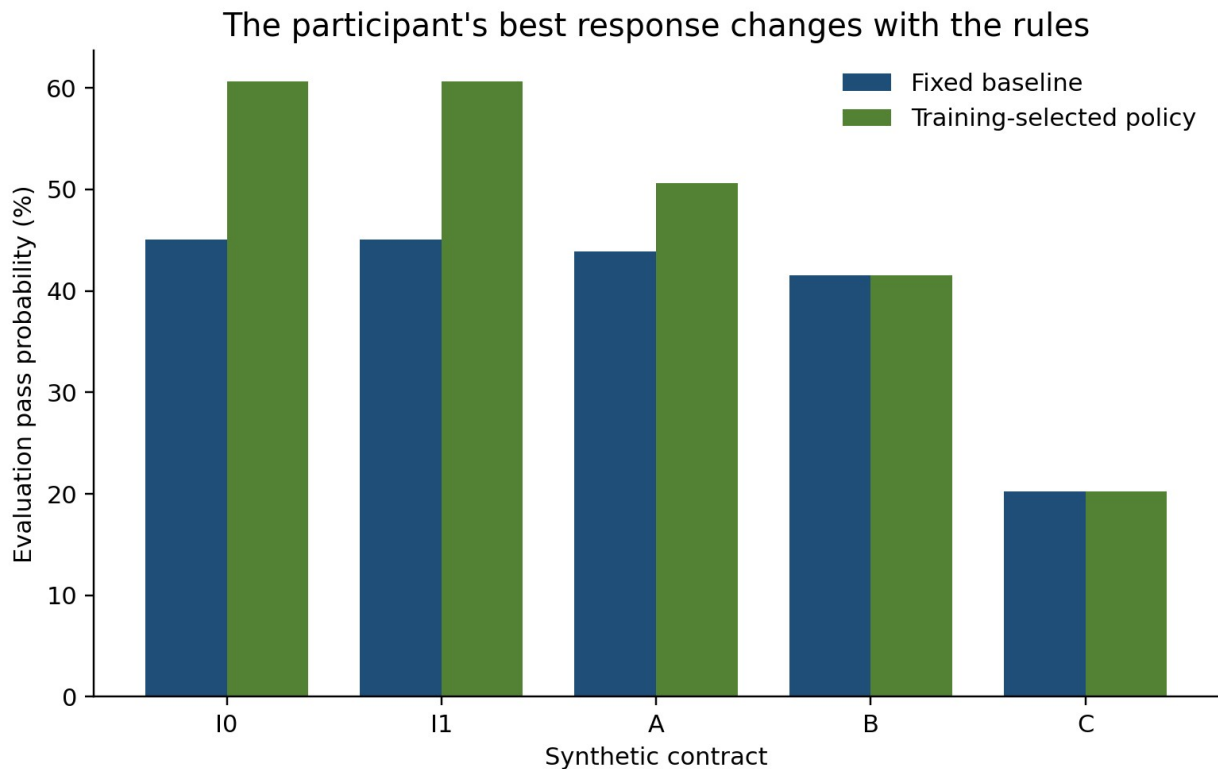


Figure 2. Fixed baseline and training-selected evaluation pass probabilities.

The static-DP transfer performs best for I0 and I1. Those systems have the same mechanical paths, so their pass probabilities must coincide. Their cash-flow values differ because I1 has a higher fee, activation, and lower payout share.

Contract A changes the selected response. Fixed high risk raises holdout pass probability from 43.90% to 50.60%. Minimum days and a 50% consistency rule do not erase the value of higher dispersion in this calibration.

Contracts B and C do. Within the tested library, fixed baseline risk remains best. The EOD trailing floor and daily-loss limit in B make the static-DP transfer less portable. C is more restrictive still: an intraday trailing floor, \$1,000 daily-loss limit, seven minimum days, and 30% consistency threshold make aggressive resolution costly.

The fact that baseline wins in B and C is not proof of a global optimum. It means only that none of the eight alternatives in the declared library improves the training objective and then survives the holdout comparison.

6.2 Contract value

Table 4. Participant expected cash flow on holdout paths

Contract	Fixed baseline EV	EV-selected policy	Selected EV
I0	\$570.74	Static-DP transfer	\$1,063.10
I1	\$414.61	Static-DP transfer	\$842.15
A	\$214.16	Fixed high	\$322.03
B	\$178.33	Fixed baseline	\$178.33
C	-\$54.39	Fixed baseline	-\$54.39

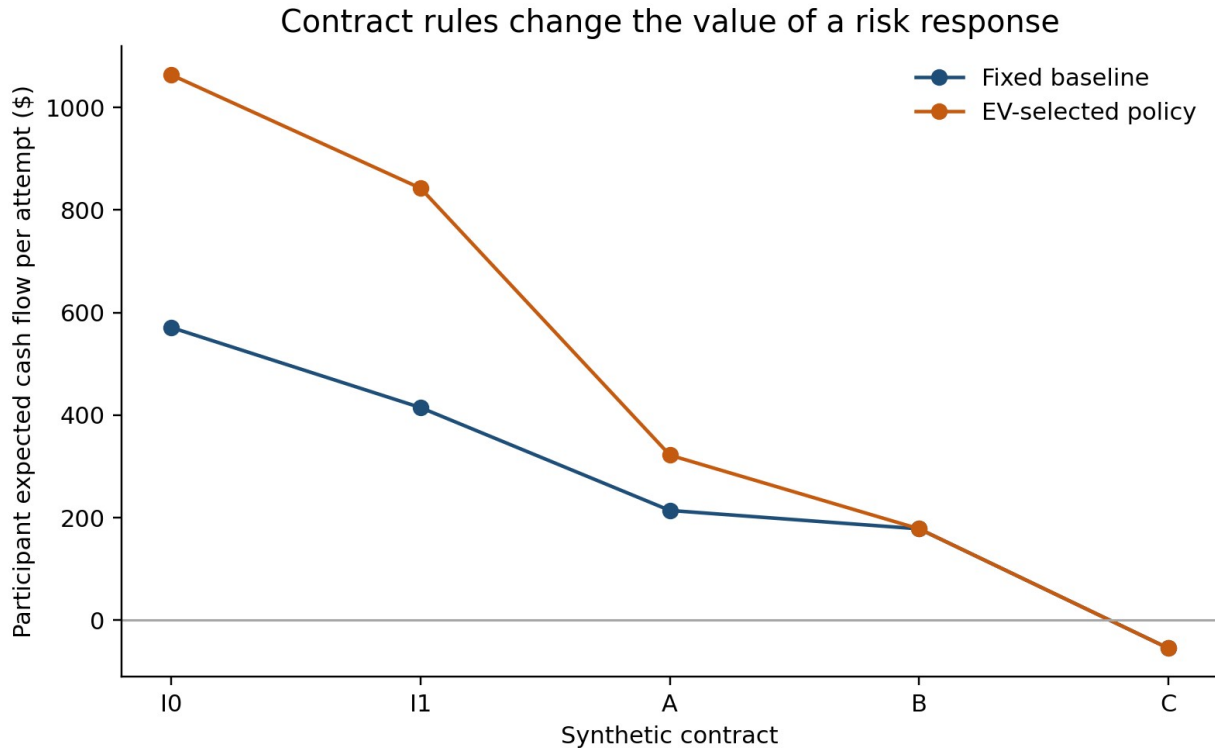


Figure 3. Participant expected cash flow under fixed baseline and the EV-selected policy.

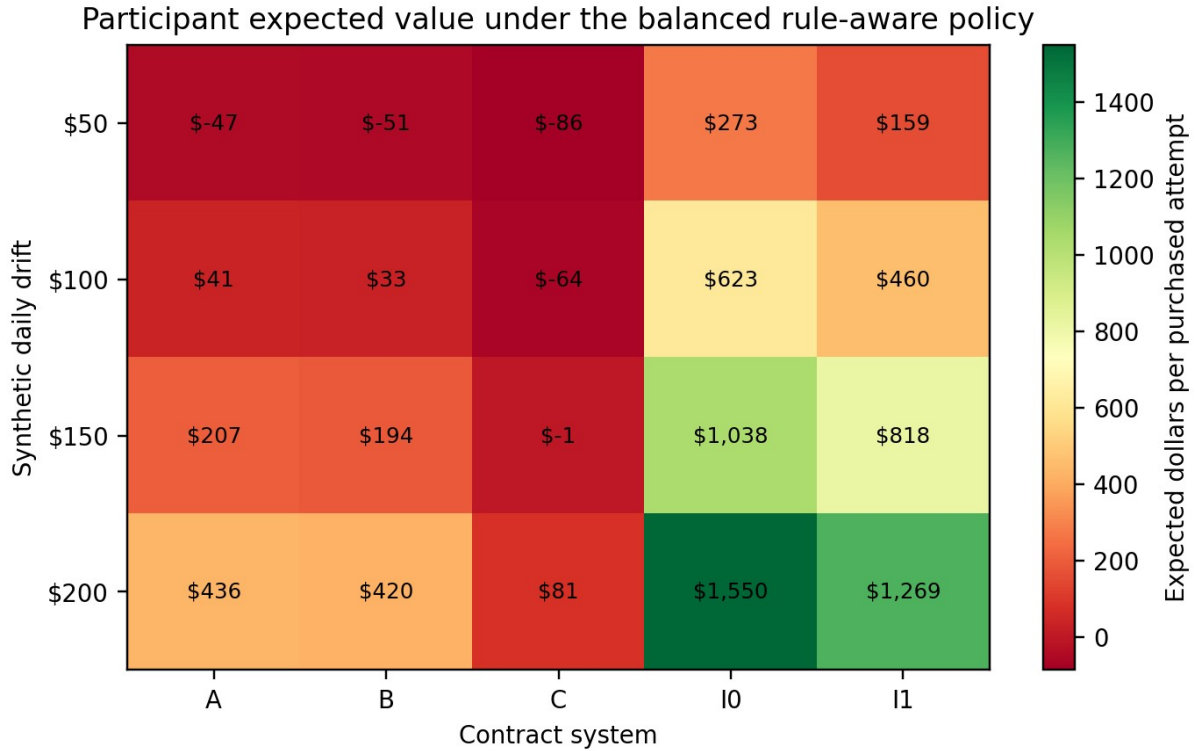
In this calibration, the pass-selected and EV-selected labels coincide for all five contracts. That is a result, not an identity. A high pass rate can fail to maximize value if the policy changes funded success, payout amount, or activation incidence differently. The workbook keeps both objectives separate so later revisions can detect such a divergence.

The large positive values for I0 and I1 should not be read as market forecasts. They come from a known positive drift, costless risk scaling, identical stage dynamics, and a large modeled payout. Contract C is the only system with negative participant EV for this favorable trader and policy library.

6.3 Expanded sensitivity surface

The analysis reruns four prespecified policies under daily drifts of \$50, \$100, \$150 and \$200 for each of the five contract systems. The resulting 80 cells use fresh common random numbers within each drift panel and 1,000 paths per stage and cell. They ask whether the policy ordering is local to the baseline drift rather than declaring a universal optimum.

The surface shows that both the best policy and the economic distance between policies depend on the contract. A policy that raises the probability of passing can lower participant value if it also raises breach risk, activation spending or the chance of reaching a funded state with little remaining loss capacity. For this reason the paper reports pass probability and expected cash flow together.



Expected-value sensitivity across drift and contract systems.

Figure 4. Participant expected value under the balanced rule-aware policy across the drift and contract surface.

6.4 Fresh holdout replications

The balanced rule-aware policy is evaluated in 20 independent holdout replications for each contract, producing 100 contract-replication observations with 1,500 paths per stage. These runs are not reused for policy selection.

Table 5. Participant expected value across fresh holdout replications

Contract	Policy	Replications	Mean EV	P10	P90
A	Rule-aware balanced	20	\$40.78	\$32.25	\$47.80
B	Rule-aware balanced	20	\$32.87	\$25.05	\$40.59
C	Rule-aware balanced	20	\$-66.07	\$-69.35	\$-62.08
I0	Rule-aware balanced	20	\$613.90	\$605.18	\$634.30
I1	Rule-aware balanced	20	\$451.65	\$444.43	\$469.30

The dispersion across replications is sampling uncertainty under the synthetic process, not uncertainty about the true distribution of trader returns. Out-of-sample robustness therefore supports the computational comparison while leaving external validity unresolved.

7. Why the Selection, Risk-Control and Robustness Implementations Report Different Pass Rates

The three papers did not originally use the same stopping convention. “Twenty sessions, \$3,000 target, and \$2,000 drawdown” is not a complete numerical definition unless the observation grid is also specified.

Table 6. Monitoring crosswalk for the Gaussian P-LV benchmark

Design	Target check	Loss check	Fixed-baseline pass
This paper: daily benchmark	Session close	Session close	45.56%
Common rule engine	Session close	Ten intraday points	45.06%
Matilla Serrano (2026c): reported benchmark	Ten intraday points	Ten intraday points	48.75%

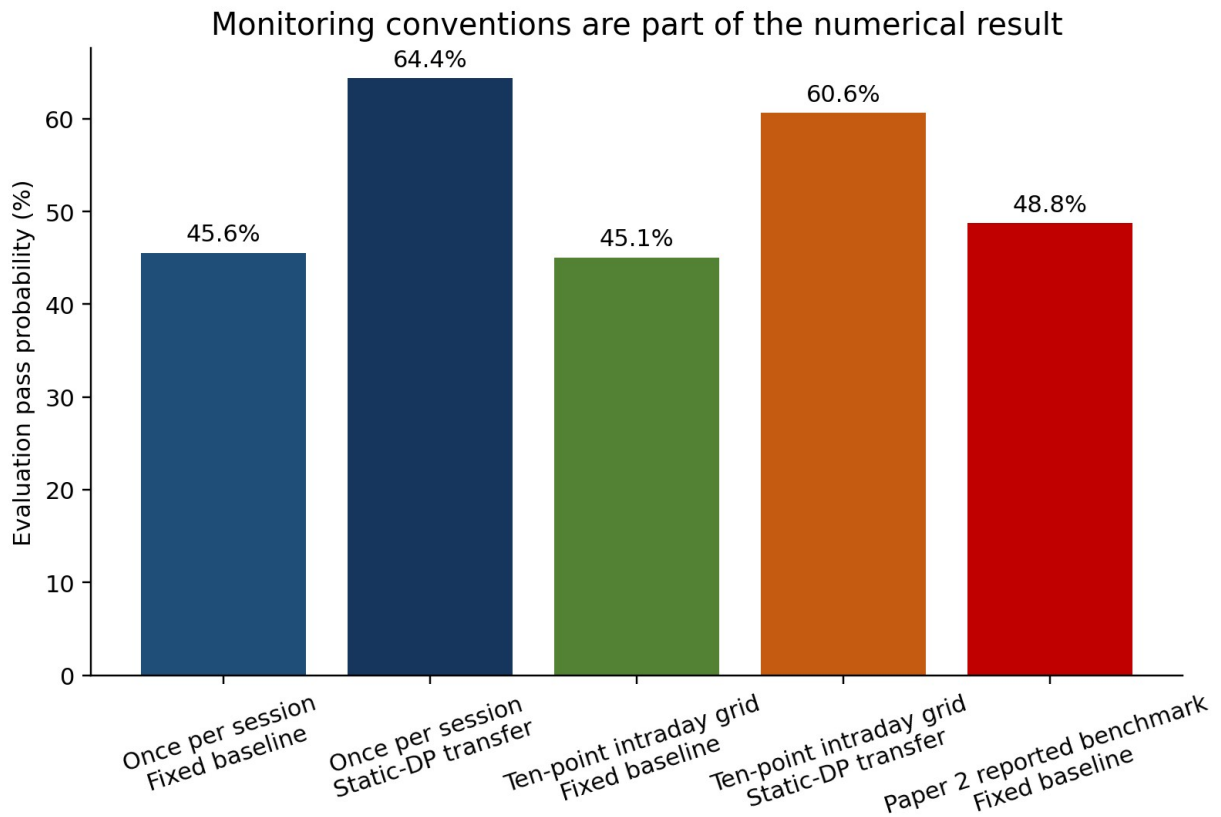


Figure 5. Monitoring crosswalk across the benchmark implementations.

Matilla Serrano (2026c) was designed as a discretized first-passage model. Both target and loss were checked on the intraday grid. A path could touch the target during the day and pass even if it later reversed before the close.

The original dynamic program in this paper was designed as a session-level control problem. The state and action changed once per session, and both barriers were checked at the resulting session close. Intraday touches were invisible.

Matilla Serrano (2026d) was designed as a rule engine. Daily loss and intraday trailing drawdown require intraday checks, but evaluation qualification was assessed after the session so minimum-day and consistency conditions could use the completed day. The revised model in this paper inherits that engine for A-C.

These choices were made for different questions: Matilla Serrano (2026c) emphasized selection under a fine first-passage approximation; this paper emphasized a manageable control state; Matilla Serrano (2026d) emphasized contract-rule sequencing. The numbers are conditional on those designs. For a direct numerical comparison, the same target check, loss check, action frequency, seed, and event ordering must be used.

8. Interpretation

First, rules are incentives. The same trader does not have one contract-independent “optimal risk.” The static contract rewards aggressive resolution near the deadline. A tight trailing and daily-loss bundle can punish the same exposure.

Second, unfinished paths matter. Low risk can reduce failure and still be poor for a finite-horizon pass objective because it leaves too many paths below target at day 20.

Third, pass probability and participant value must remain separate. This revision reports both, even though they select the same policies in the baseline calibration.

Fourth, an optimal label must state its domain. The static DP is optimal only on its declared grid, action set, process, monitoring convention, and objective. The A-C result is a validated finite-policy comparison.

9. Limitations

Drift and volatility are known, constant, and scaled linearly by exposure. The model ignores estimation error, changing opportunities, skewness, jumps, serial dependence, commissions, slippage, liquidity, and behavioral execution. A separate robustness analysis examines several synthetic process alternatives but does not solve the full control problem under each.

Risk changes once per day. A participant who can react intraday faces a different control problem. The policy library is transparent but sparse; a richer search could improve A-C outcomes or overfit the simulation.

The rule systems are synthetic archetypes and do not identify a named provider. The 20-session horizon applies to both stages for comparability. No later payouts, scaling, account recycling, or capital constraints are modeled.

10. Conclusion

Matilla Serrano (2026d) showed that changing the rules changes outcomes. This paper shows the other half: changing the rules changes the participant’s best tested response. The static dynamic-programming transfer is selected under I0 and I1, fixed high risk under A, and fixed baseline under B and C.

For the favorable P-LV trader, the selected holdout pass probability falls from 60.64% in the ideal systems to 20.26% in Contract C. The result cannot be summarized by the common Target-to-Drawdown Ratio. Minimum days, consistency, daily loss, trailing frequency, payout buffer, and cap alter both the probability map and the response to it.

Monitoring conventions also matter. The 48.75% benchmark in Matilla Serrano (2026c) and the 45.56% daily benchmark in this paper answer different stopping-time questions. Once observation timing is made explicit, the difference is explainable rather than contradictory.

References

- Basak, S., Pavlova, A., and Shapiro, A. (2007). Optimal asset allocation and risk shifting in money management. *Review of Financial Studies*, 20(5), 1583-1621. <https://doi.org/10.1093/rfs/hhm026>
- Browne, S. (1999). Reaching goals by a deadline: Digital options and continuous-time active portfolio management. *Advances in Applied Probability*, 31(2), 551-577. <https://doi.org/10.1239/aap/1029955147>
- Carpenter, J. N. (2000). Does option compensation increase managerial risk appetite? *Journal of Finance*, 55(5), 2311-2331. <https://doi.org/10.1111/0022-1082.00288>
- Chekhlov, A., Uryasev, S., and Zabarankin, M. (2005). Drawdown measure in portfolio optimization. *International Journal of Theoretical and Applied Finance*, 8(1), 13-58. <https://doi.org/10.1142/S0219024905002767>
- Cvitanic, J., and Karatzas, I. (1995). On portfolio optimization under drawdown constraints. In *Mathematical Finance*, IMA Volumes in Mathematics and its Applications, 65, 35-45. https://doi.org/10.1007/978-1-4757-2435-6_3
- 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. (2026b). *Market Edge versus Contract Edge: Negative-Expectancy Trading under an Idealized Evaluation Contract*. SSRN preprint. <https://doi.org/10.2139/ssrn.7468080>.
- Matilla Serrano, F. (2026c). *Skill or Variance? What Trading Evaluations Actually Select*. SSRN preprint. <https://doi.org/10.2139/ssrn.7473404>.
- Matilla Serrano, F. (2026d). *The Rules Are the Model: Path Dependence, Payout Frictions, and Selection in Synthetic Trading-Evaluation Contracts*. SSRN preprint. <https://doi.org/10.2139/ssrn.7479818>.
- Nystrup, P., Boyd, S., Lindström, E., and Madsen, H. (2019). Multi-period portfolio selection with drawdown control. *Annals of Operations Research*, 282(1-2), 245-271. <https://doi.org/10.1007/s10479-018-2947-3>
- 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>

Appendix A. Interpretation of “Best”

For I0 and I1, the daily static DP remains the numerical optimum for its finite model. When transferred to the ten-point rule engine it becomes a candidate policy. For A-C, “selected” or “best response” always means the highest training value among the nine declared policies, subsequently reported on holdout paths. It does not mean a global solution to the full path-dependent control problem.

Appendix B. Objectives, policy comparisons, and feasible implementation

B.1 Passing is not the terminal economic objective

Let $p(a)$ be the probability of an evaluation pass under policy a , and let $G(a)$ be the expected participant receipt from a newly activated funded account under that policy. G includes a zero receipt after funded failure. The value of buying an evaluation is

$$J(a) = p(a)[G(a) - A] - F.$$

For two policies a and b , an exact decomposition is

$$J(a) - J(b) = [p(a) - p(b)][G(b) - A] + p(a)[G(a) - G(b)].$$

The first term values the change in access to the funded stage at the old net continuation value. The second values the change in the funded opportunity itself. A policy can therefore increase passing but reduce value: more frequent activation is costly when the funded opportunity is weak, and a pass-oriented exposure rule may damage subsequent payout eligibility. The numerical coincidence of the selected policies in Table 4 does not eliminate this distinction.

The entry fee F cancels in a comparison after the evaluation has been purchased. It remains essential to the prior decision whether to buy at all. Re-entry would create a separate lifecycle decision problem rather than justify treating repeated fees as sunk before they are paid.

B.2 Measuring improvement without selecting the winner twice

On common holdout paths, define a paired difference d_i between the selected policy's success indicator and the fixed benchmark's indicator. The estimated improvement is the average of d_i . Its uncertainty depends on the variation of these differences, not on adding two unrelated binomial standard errors. Common paths can substantially reduce the noise of a policy comparison when both policies react similarly to market shocks.

The original simulations use antithetic paths. For an inferential comparison, the independent sampling unit is an antithetic pair: first average the two differences within each pair, then calculate the standard error across pairs. Treating both members as independent would ignore their designed covariance. No new standard errors are claimed in this appendix because the existing tables report point estimates. A future uncertainty table should use retained pair-level outputs or regenerate them, not infer a covariance from rounded probabilities.

If multiple policies tie on training results, a deterministic tie-breaking rule should be fixed in advance. Looking at the holdout and then choosing whichever policy happens to win converts the holdout into a second training set. Further refinement is possible, but it needs a new evaluation sample or a clearly labeled exploratory analysis.

B.3 What a feasible risk policy must specify

A multiplier is not yet an executable order. An implementation would need the instrument, contract granularity, point value, commissions, margin and position limits, and the treatment of a signal too small to trade. Multiplying both drift and volatility by the same a assumes scalable exposure without a change in signal quality or execution cost. It is useful for a controlled experiment but not a guarantee that real trading scales linearly.

The timing of information also matters. A daily policy may use only information available when its action is chosen. Session-end profit, future volatility and tomorrow's rule state cannot be inputs to today's multiplier. A path-dependent solver must carry sufficient state for the current

floor, high-water mark, consistency statistic and elapsed days; current P&L alone is not generally Markov for the contractual payoff.

The next computational extension should compare a larger feasible policy library on the same five systems, reporting the extra improvement against its additional search cost. A small positive gain would be worth reporting without relabeling it a global optimum. Conversely, a failure to improve the benchmark would be informative about this library, not proof that no better strategy exists.

Declarations

Funding. No external funding was received for this research.

Conflicts of interest. No financial or commercial ties exist with any trading-evaluation provider.

Preprint DOI. [SSRN DOI pending].

Supplementary materials. The workbook, source code, and documented inputs are available in the Zenodo replication package:

<https://doi.org/10.5281/zenodo.22770675>

Advice boundary. The paper is academic analysis, not investment or trading advice.