

TRAVIDENCE

D2 — Executive Brief

Audit-Grade Pre-Validation of Trading Strategies — Executive Brief

Ignacio Arias · Travidence, LLC

ORCID 0009-0001-7503-3830

Version 1.0 — August 2026

The audit-grade alternative for prop firm challengers.

§1. The 86% problem

A prop firm challenge on a standard 50,000–100,000 USD account typically costs between 300 and 600 USD per attempt; smaller accounts start at a few tens of dollars. At retail prop firms the cycle is the same: the trader pays the fee, attempts the evaluation under the firm’s rules (return target, trailing drawdown limit, minimum trading days), and the firm decides whether to issue the funded account. The commercial argument is direct: if your strategy has edge, the firm absorbs the capital risk and shares the upside with you.

Aggregate FPFx Tech data on more than 300,000 accounts, reported by Finance Magnates (2024), describes what happens next with uncomfortable precision:

- **86%** fail the initial evaluation.
- Of the 14% who pass, only **7%** of the total reach a payout — the rest pass but never reach a withdrawal from the funded account, typically through a rules breach (trailing drawdown, maximum daily loss, directional exposure).
- The remaining **93%** pay the fee and commonly restart the cycle.

The pattern is not statistical noise or individual bad luck. It is systematic evidence that most challengers enter the desk with strategies whose in-sample metrics are products of overfitting — the technical phenomenon the academic literature documents as *backtest overfitting* (Bailey, Borwein, López de Prado, and Zhu, 2017). The trader optimizes parameters over the history until the equity curve looks acceptable, and the result gets reported as “edge”; what was actually captured is the best combination of noise over the historical sample, not a persistent mechanism.

Expected value for the average challenger is negative before paying the fee. With a 14% probability of passing the challenge and a 50% conditional probability of not breaking the funded account (that is, a 7% absolute payout rate), the expected value of a typical challenge sits several hundred dollars below the fee paid. Failed challengers commonly attempt again — paying a second or third fee on the same strategy with the same structural problem — turning the expected loss into a compounding pattern across attempts. The operational question is direct: is there a way to evaluate the strategy *before* paying the challenge, and to identify which attempts are economically irrational from the start?

§2. Why existing tools do not solve the problem

Four classes of actor partially occupy the quantitative validation space. None provides an independent service at audit grade.

Category	Structural limitation
Implementation libraries (mlfinlab, RiskLabAI)	High-quality technical infrastructure that implements the López de Prado stack in code, but does not issue certification. The analyst interprets and reports the metrics; no third party signs the verdict and no pre-specified threshold triggers a PASS/FAIL binary decision for commercial use.
Retail consumer validators	Aggregate score on the [0,100] scale without hard gates. A strategy with severe regime asymmetry can accumulate a numerically acceptable score because favorable dimensions compensate for unfavorable ones. The system does not distinguish continuous weakness from binary structural failure.
The prop trading firms themselves	An inescapable structural conflict of interest. Their model depends on evaluation-fee flow; effective pre-validation would shrink the challenger universe — and with it the firm's revenue. A firm that sells challenge slots cannot simultaneously warn the client that the strategy is not viable.
Trade journals and replay tools	Closed-operations tracking and psychological journaling tools. Useful for personal discipline but they do not validate whether the strategy has genuine edge or whether the in-sample metrics are statistical artifacts.

The specific gap is the combination of four attributes: (i) serious technical implementation of the López de Prado stack at the level of the implementation libraries, (ii) a service layer with a verdict signed by an independent third party, (iii) a hard-gate system that distinguishes cliff-edge failure (regime asymmetry, monotonic decay, selection overfitting) from continuous weakness, and (iv) a public result band that speaks simultaneously to the retail challenger and to professional stakeholders (reviewers, B2B counterparties, future scrutiny).

That strip, in other regulated markets occupied by independent audit firms, has no equivalent incumbent in retail strategy pre-validation.

§3. Introducing TRAVIDENCE

The Big Four firms (Deloitte, EY, PwC, KPMG) operate in the financial assurance market as an independent third party. They sign opinions — Unqualified, Qualified, Adverse, Disclaimer — on financial statements under codified procedural standards (GAAP/IFRS) and auditing standards

(SAS/ISA). Their market value derives from three converging attributes: independence (they do not provide consulting that creates conflict with the verdict), standards (procedural discipline is documented and retrospectively auditable), and reputation (the year-on-year verdict track record sustains the service premium).

TRAVIDENCE positions itself as the conceptual equivalent in the quantitative strategy assurance market, with three contextual adjustments: the audited object is the strategy's validation methodology (not financial statements), the procedural standards are the TRAVIDENCE methodology documented in the D1 Technical Paper, and the retail-plus-B2B-tier-2 target market is structurally underserved by the Big Four incumbents. The TRAVIDENCE verdict does not claim functional equivalence to a financial audit—it is not an opinion on financial statements under GAAP/IFRS—but it does deliberately adopt that procedural discipline as a base of legal and reputational defensibility.

Audit-grade independent validation, before you pay the challenge fee.

TRAVIDENCE is built for challengers who take their trading seriously: traders returning after a failed challenge, quants with technical backgrounds, and prop firms that need institutional pre-screening. This is not a generic retail tool—it is audit-grade pre-validation that reduces the material risk of your next challenge.

The distinction matters. A retail consumer analysis processes a CSV, computes an aggregate score, and emits a classification; the full cycle takes less than a minute. A TRAVIDENCE verdict uses the same technical stack as the institutional implementation libraries plus five proprietary methodological contributions, runs it against documented procedural standards, applies hard gates that halt certification when one dimension exhibits cliff-edge failure, and emits a signed verdict with a public result band, inspired by the procedural discipline of AICPA SAS-122 / ISA 705. The difference is not compute time but the standard of the output: a verdict defensible under professional review—not a consumer score.

The five validators in retail language

Validator	What it detects
Deflated Sharpe Ratio (DSR)	Whether your Sharpe Ratio looks good only because you tried many configurations until one worked. Penalizes the expected Sharpe inflation when the best of N attempts is reported.
Capital-normalized Calmar	Drawdown measured against real capital at risk, not against broker margin. Strips the leverage-induced inflation that distorts cross-asset comparisons.
Sub-period monotonic decay	Whether your edge is dying over time. Splits the backtest period into contiguous chunks and evaluates whether the metric decays systematically from the first to the last.

Validator	What it detects
Component isolation	For multi-component strategies, verifies that each component contributes real edge. Detects aggregates whose healthy Sharpe depends on combination logic rather than on each signal's actual edge.
MUST-PASS regime	Tests survival in both high-volatility and low-volatility regimes. A strategy that survives in only one regime will meet its weak side when the market shifts.

The Robustness Score

A score on the 0–100 scale that integrates the outputs of the five validators into a composite metric. The structure is deliberately hybrid: a continuous weighted aggregation of three quality dimensions (Edge Quality, Risk Profile, Sample & Coverage) combined with a hard-gate system that acts as caps when a dimension exhibits structural failure.

The score translates into a public band the challenger can read directly: **Robust, Conditional Robustness, Limited Robustness, Not Robust, Disclaimer of Opinion**. The same band is traceable to the institutional criterion that produced it — the information a professional reviewer, a B2B counterparty, or future media scrutiny needs. The Big Four parallel is deliberate: TRAVIDENCE adopts the vocabulary and procedural discipline of the professional auditing body as a base of legal and reputational defensibility.

§4. How it works

Each of the five validators attacks a specific failure mode documented in the academic literature and operationalized by TRAVIDENCE as a verdict invariant.

Deflated Sharpe Ratio with dual n_trials reporting. The standard Sharpe Ratio does not penalize the number of hypotheses explored. If you tested 200 variants of your strategy and reported the best, your Sharpe is inflated in a predictable way — not because you lied but because the best of 200 samples is, by statistical construction, biased upward relative to the true mean. The TRAVIDENCE contribution is to report two parallel versions: `dsr_strict` with the conservative institutional count (every variant that ever touched data, including abandoned ones) and `dsr_session` with the comfortable current-session count. The gap between the two makes the selection bias auditable — bias that single reporting hides.

Capital-normalized Calmar. The standard retail Calmar — annualized return divided by maximum percent drawdown — uses the rolling peak of the equity curve as denominator. In leveraged contexts (futures, retail FX under ESMA 30:1, crypto perpetual with cross-margin), the rolling peak grows with the controlled notional, not with the capital actually at risk. The result is a Calmar inflated in proportion to the leverage embedded in margin sizing. TRAVIDENCE reformu-

lates against `capital_baseline` defined as `nominal_capital / leverage_factor` — the trader’s real margin posted, the actual skin in the game. The resulting metric (`Calmar_capital_frame`) is comparable across asset classes and strips the typical retail illusion of “my Calmar is 2.5.”

Sub-period monotonic decay. Aggregate metrics over the full history hide temporal trajectories. A strategy with aggregate profit factor of 1.8 may conceal a fall from $PF = 2.5$ in the earliest period to $PF = 1.1$ in the most recent — the arithmetic mean does not capture that fact. TRAVIDENCE splits the return series into K contiguous chunks of equal length (default five), evaluates a configurable metric (Sharpe by default) on each chunk, and emits a binary verdict on the trajectory. The rule is strict: the sequence must not be monotonically non-increasing *and* the endpoint-to-endpoint loss must be below the threshold. Both conditions simultaneously, not either.

Component isolation. Multi-component strategies — ensembles of two or more independent signals combined in a signal aggregator or portfolio — can produce an aggregate with a healthy Sharpe by combination noise, with no individual component having edge statistically distinguishable from noise. Profit factor, Calmar, and DSR applied to the aggregate may all look healthy; yet the fragility is structural, the edge is not replicable, and it does not survive minor changes in combination logic. TRAVIDENCE decomposes the aggregate Sharpe into the sum of standalone per-component Sharpes plus an interaction residual. If the residual exceeds the arithmetic floor of the problem, the system flags portfolio-level overfitting.

MUST-PASS regime. A strategy that performs well in low-volatility regimes and catastrophically in high-volatility regimes may accumulate an acceptable aggregate Sharpe if the time fraction in each regime is functional to the backtest’s directional bias. Silent cross-regime averaging is invisible to the client without an explicit breakdown. TRAVIDENCE labels each observation as vol-on (high annualized volatility, per per-asset-class threshold) or vol-off (low volatility), evaluates the strategy separately in both regimes, and demands that it clear the minimum criteria in *both*. The rule is AND, not average. A failure in a single regime triggers the system’s most severe gate.

Robustness Score band table

Score	Band	Recommended action
80–100	Robust	The strategy is ready for the prop firm challenge. The TRAVIDENCE seal applies.
65–79	Conditional Robustness	Candidate profile with specific weaknesses listed in the report. Refine and resubmit.
50–64	Limited Robustness	Deficiencies detected. Structural changes are required before the challenge.
0–49	Not Robust	Strategy not viable in its current form. Likely overfitting, regime asymmetry, or decay. Significant changes are recommended.

Score	Band	Recommended action
N/A	Disclaimer of Opinion	Insufficient data to certify. Generate more trades or extend calendar coverage; resubmit.

Technical ↔ retail translation

Technical term	What it means for the challenger
Selection bias	The Sharpe you report is the best of many tries, not the representative one
Capital frame	Your drawdown measured against what you actually risk, not against what the broker lets you control
Monotonic decay	Your strategy no longer works the way it used to — the edge is dying
Interaction residual	Your aggregate depends on combination logic, not on each signal's real edge
Regime asymmetry	When the market shifts, your strategy will meet its weak side

§5. Three real cases, three outcomes

TRAVIDENCE applied the methodology to three textbook strategies on public data. The cases were chosen to span different evidence situations — not to demonstrate that every strategy fails, but to illustrate how the framework discriminates: it refuses to score when the evidentiary base is insufficient, issues a scored verdict when the evidence warrants one, and gates on structural failure when the evidence supports that conclusion.

Case 01 — Golden Cross on CL crude (2000–2025). SMA(50) / SMA(200) crossover on WTI crude futures. Aggregate Sharpe 0.32, profit factor 3.26, maximum drawdown –62.2%, 18 trades over 25 years. The trade count falls below the admissibility floor; validators run and produce metrics — what is withheld is the composite score (Disclaimer of Opinion for sample floor). **Verdict: Disclaimer of Opinion.** A professional refusal to opine, not a failure classification. The mechanism may be sound; the evidentiary base cannot support a defensible verdict under institutional standards.

Case 02 — RSI(14) mean reversion on BTC perpetual (2020–2025). Long entry when RSI<30, exit when RSI>50. Sharpe 0.41, profit factor 1.96, maximum drawdown –47.9%, 31 trades over 5 years. Trade count again below the admissibility floor; validators run and produce metrics — what is withheld is the composite score (Disclaimer of Opinion for sample floor). **Verdict: Disclaimer of Opinion.** Same structural reason as Case 01. Resubmit with an extended sample.

Case 03 — Asian Range Breakout on EUR/USD (2015–2025) — the original audit’s anchor case. Breakout of the Asian session range during the London open. Sharpe 2.67, profit factor 1.76, maximum drawdown -6.5%, 2,584 trades over 10 years — numbers that in isolation would suggest a clear high-band profile. DSR clears with margin (1.0), capital-frame Calmar clears the institutional threshold, the sub-period does not decay, and both volatility regimes clear the must-pass criteria. No gate is triggered; the strategy receives a public Robustness Score — the first scored case in this set. **Verdict: Limited Robustness (56/100).** The score lands below the Robust band, reflecting structural deficiencies identified across the quality dimensions. The certificate discloses each dimension requiring attention before a challenge attempt.

A retail aggregation tool would compress all three cases into a single continuous ranking. TRAVIDENCE’s verdict architecture distinguishes them by output type: Disclaimer of Opinion when the evidentiary base does not reach the institutional floor, a Limited Robustness score when evidence is present but quality falls short of the Robust band. No textbook strategy earns the Robust seal — not because the framework is adversarial, but because these are strategies whose published record does not meet the standard.

- **Figure 1.** TRAVIDENCE certificate — Case 01 — Path: `../../../../case_studies/01-golden-cross-cl/audit_certificates/01-cl-golden-cross-audit-certificate-en.pdf`

TRAVIDENCE

Independent strategy validation

STRATEGY	CL Crude Oil Golden Cross (50/200 SMA)
ASSET CLASS	Commodity Futures — CL Crude Oil
PERIOD COVERED	2000-08-23 to 2025-05-16
CERTIFICATE ID	TRV-E8A6-E8E8
SCORED UNDER	Robustness Score v1.1.0 (2026-Q3)
ISSUED	2026-08-17 04:42 UTC

ROBUSTNESS SCORE

N/A

Disclaimer of Opinion

VERDICT

This strategy cannot be certified under TRAVIDENCE standards: n_trades = 18 below the admissibility floor Generate additional trade data and resubmit.

GATE STATUS

GATE	STATUS	EFFECT
Regime Must-Pass	PASSED	—
Subperiod Monotonic Decay	PASSED	—
Component Isolation	PASSED	—
Combined	Final outcome: no gates triggered	

TRAVIDENCE

Independent strategy validation

VALIDATOR DETAIL

Deflated Sharpe Ratio

DSR variant	DSR	n_trials	E[max SR] (annualized)
Strict	0.964	2	0.111
Session	0.964	2	0.111

Capital-Normalized Calmar

Frame	Calmar	Notes
Margin frame	0.295	baseline = 5,555.56 (leverage 9x)
Capital frame	0.105	inflation ratio = 2.808

Subperiod Monotonic Decay

Field	Value
Status (no decay)	FAIL
Metric	sharpe
Per-subperiod (5 chunks)	0.57, 0.54, -0.72, -0.42, 0.37
Max consecutive drop	234.59%
Overall decay	34.49%

Regime Must-Pass

Regime	Status	min Sharpe	max DD	n obs
vol_off	FAIL	nan	0.00%	5205
vol_on	FAIL	1.130	17.36%	338

Component Isolation: N/A — single-component strategy

METHODOLOGY REFERENCES

Bailey, D. H., & López de Prado, M. (2014). *The Deflated Sharpe Ratio: Correcting for Selection Bias, Backtest Overfitting, and Non-Normality*. Journal of Portfolio Management 40(5).

Lo, A. W. (2002). *The Statistics of Sharpe Ratios*. Financial Analysts Journal 58(4): 36-52.

Young, T. W. (1991). *Calmar Ratio: A Smoother Tool*. Futures Magazine, October.

AICPA SAS-122 / IAASB ISA 705 & 706. *Modifications to the Opinion in the Independent Auditor's Report*.

Reproducibility Hash (full):

sha256:e8a6e8e8fd238f82fa65da2c9eb2d20f1e8a9d8635b489e8ad9c752966d1de8c · hash schema 2

v1.1.0 sigma limitation note: This release uses conservative zero-sigma handling for confidence interval propagation. Single-realization Sharpe and Calmar inputs have sigma=0 by convention; full propagation through the delta method is deferred to a future release. CIs shown reflect this conservative bound; tighter intervals require bootstrap or block-bootstrap resampling.

Calibration disclosure: This certificate was issued under v1.1.0 (2026-Q3), which represents a best-achievable calibration of the Robustness Score scoring system. The calibration is conservative by design: strategies receiving Not Robust may include some that could be classified Limited Robustness under refined calibration. Subsequent recalibration with production data is planned as a quarterly protocol.

Generated by the TRAVIDENCE audit engine. Scored under v1.1.0 | calibration 2026-Q3.

Issued by Travidence, LLC · Wilmington, Delaware, USA

Travidence, LLC maintains technology professional liability and cyber (Tech E&O/Cyber) insurance in force.

- **Figure 2.** TRAVIDENCE certificate — Case 02 — Path: ../../../../case_studies/02-rsi-mean-reversion-btc/audit_certificates/02-btc-rsi-mean-reversion-audit-certificate-en.pdf

STRATEGY	BTC Perpetual RSI 14 Mean Reversion
ASSET CLASS	Crypto Perpetual — BTC/USDT
PERIOD COVERED	2020-01-02 to 2025-05-19
CERTIFICATE ID	TRV-6A8E-058C
SCORED UNDER	Robustness Score v1.1.0 (2026-Q3)
ISSUED	2026-08-17 04:42 UTC

ROBUSTNESS SCORE

N/A

Disclaimer of Opinion

VERDICT

This strategy cannot be certified under TRAVIDENCE standards: n_trades = 31 below the admissibility floor Generate additional trade data and resubmit.

GATE STATUS

GATE	STATUS	EFFECT
Regime Must-Pass	PASSED	—
Subperiod Monotonic Decay	PASSED	—
Component Isolation	PASSED	—
Combined	Final outcome: no gates triggered	

TRAVIDENCE

Independent strategy validation

VALIDATOR DETAIL

Deflated Sharpe Ratio

DSR variant	DSR	n_trials	E[max SR] (annualized)
Strict	0.691	3	0.380
Session	0.691	3	0.380

Capital-Normalized Calmar

Frame	Calmar	Notes
Margin frame	0.789	baseline = 1,500.00 (leverage 20x)
Capital frame	0.572	inflation ratio = 1.379

Subperiod Monotonic Decay

Field	Value
Status (no decay)	PASS
Metric	sharpe
Per-subperiod (5 chunks)	0.39, 0.98, -0.12, 1.01, 1.71
Max consecutive drop	112.66%
Overall decay	-332.04%

Regime Must-Pass

Regime	Status	min Sharpe	max DD	n obs
vol_off	FAIL	nan	0.00%	1228
vol_on	PASS	1.569	14.67%	596

Component Isolation: N/A — single-component strategy

METHODOLOGY REFERENCES

Bailey, D. H., & López de Prado, M. (2014). *The Deflated Sharpe Ratio: Correcting for Selection Bias, Backtest Overfitting, and Non-Normality*. Journal of Portfolio Management 40(5).

Lo, A. W. (2002). *The Statistics of Sharpe Ratios*. Financial Analysts Journal 58(4): 36-52.

Young, T. W. (1991). *Calmar Ratio: A Smoother Tool*. Futures Magazine, October.

AICPA SAS-122 / IAASB ISA 705 & 706. *Modifications to the Opinion in the Independent Auditor's Report*.

Reproducibility Hash (full):

sha256:6a8e058c61179010c5be84e68f846e758acd4de1453d37b9b1539b317b5babcb6 · hash schema 2

v1.1.0 sigma limitation note: This release uses conservative zero-sigma handling for confidence interval propagation. Single-realization Sharpe and Calmar inputs have sigma=0 by convention; full propagation through the delta method is deferred to a future release. CIs shown reflect this conservative bound; tighter intervals require bootstrap or block-bootstrap resampling.

Calibration disclosure: This certificate was issued under v1.1.0 (2026-Q3), which represents a best-achievable calibration of the Robustness Score scoring system. The calibration is conservative by design: strategies receiving Not Robust may include some that could be classified Limited Robustness under refined calibration. Subsequent recalibration with production data is planned as a quarterly protocol.

Generated by the TRAVIDENCE audit engine. Scored under v1.1.0 | calibration 2026-Q3.

Issued by Travidence, LLC · Wilmington, Delaware, USA

Travidence, LLC maintains technology professional liability and cyber (Tech E&O/Cyber) insurance in force.

- **Figure 3.** TRAVIDENCE certificate — Case 03 — Path: ../../../../case_studies/03-asian-range-breakout-eurusd/audit_certificates/03-eurusd-asian-range-breakout-audit-certificate-en.pdf

STRATEGY	EUR/USD Asian Range Breakout
ASSET CLASS	Forex Major — EUR/USD
PERIOD COVERED	2015-01-01 to 2025-05-19
CERTIFICATE ID	TRV-AEC6-AD46
SCORED UNDER	Robustness Score v1.1.0 (2026-Q3)
ISSUED	2026-08-17 04:42 UTC

ROBUSTNESS SCORE

56 / 100

Limited Robustness

VERDICT

The strategy shows continuous quality below the Robustness threshold and exhibits material weaknesses across one or more validators. Structural rework required before challenge.

GATE STATUS

GATE	STATUS	EFFECT
Regime Must-Pass	PASSED	—
Subperiod Monotonic Decay	PASSED	—
Component Isolation	PASSED	—
Combined	Final outcome: no gates triggered	

TRAVIDENCE

Independent strategy validation

VALIDATOR DETAIL

Deflated Sharpe Ratio

DSR variant	DSR	n_trials	E[max SR] (annualized)
Strict	1.000	3	0.260
Session	1.000	3	0.260

Capital-Normalized Calmar

Frame	Calmar	Notes
Margin frame	2.360	baseline = 3,333.33 (leverage 30x)
Capital frame	1.728	inflation ratio = 1.366

Subperiod Monotonic Decay

Field	Value
Status (no decay)	PASS
Metric	sharpe
Per-subperiod (5 chunks)	3.17, 2.11, 3.54, 4.46, 2.74
Max consecutive drop	38.49%
Overall decay	13.42%

Regime Must-Pass

Regime	Status	min Sharpe	max DD	n obs
vol_off	PASS	3.762	1.85%	1343
vol_on	PASS	2.782	4.83%	1343

Component Isolation: N/A — single-component strategy

METHODOLOGY REFERENCES

Bailey, D. H., & López de Prado, M. (2014). *The Deflated Sharpe Ratio: Correcting for Selection Bias, Backtest Overfitting, and Non-Normality*. Journal of Portfolio Management 40(5).

Lo, A. W. (2002). *The Statistics of Sharpe Ratios*. Financial Analysts Journal 58(4): 36-52.

Young, T. W. (1991). *Calmar Ratio: A Smoother Tool*. Futures Magazine, October.

AICPA SAS-122 / IAASB ISA 705 & 706. *Modifications to the Opinion in the Independent Auditor's Report*.

Reproducibility Hash (full):

sha256:aec6ad462f11f53609e946864136edfb953551bcca892906aa713243e1d2db5d · hash schema 2

v1.1.0 sigma limitation note: This release uses conservative zero-sigma handling for confidence interval propagation. Single-realization Sharpe and Calmar inputs have sigma=0 by convention; full propagation through the delta method is deferred to a future release. CIs shown reflect this conservative bound; tighter intervals require bootstrap or block-bootstrap resampling.

Calibration disclosure: This certificate was issued under v1.1.0 (2026-Q3), which represents a best-achievable calibration of the Robustness Score scoring system. The calibration is conservative by design: strategies receiving Not Robust may include some that could be classified Limited Robustness under refined calibration. Subsequent recalibration with production data is planned as a quarterly protocol.

Generated by the TRAVIDENCE audit engine. Scored under v1.1.0 | calibration 2026-Q3.

Issued by Travidence, LLC · Wilmington, Delaware, USA

Travidence, LLC maintains technology professional liability and cyber (Tech E&O/Cyber) insurance in force.

Note on the “Gate Status” panel (Figures 1–3): in a Disclaimer of Opinion verdict, gate evaluation never occurs — the sample-admissibility check stops the flow before any composite score exists for a gate to cap. The panel reports that non-binding state; the underlying per-validator results, two of which are not satisfied in Cases 01 and 02, are reported in the “Validator Detail” table on the same page. The certificate template lacks a third visual state and renders non-binding as “PASSED”; the same two-state limitation applies to the Component Isolation row when the test reports N/A for single-component strategies (Figures 1–3). This presentation limitation is registered for correction in a future release of the certificate template.

Three cases, three distinct verdict types. The transition from aggregate-favorable metrics to a Limited Robustness score — and from thin-sample strategies to Disclaimer — is the discrimination that defines audit-grade pre-validation. If your strategy has a similar evidence situation to any of the three, the verdict will reflect the structural reality, not an aggregate that rounds weaknesses upward. That is the 86% the FPF data documents: strategies whose structural profile predicts failure, evaluated before the challenge fee.

§6. What you get

Turnaround

Indicative turnaround: the validation pipeline is automated end-to-end — the verdict and certificate are issued when processing completes, typically within the same business day from receipt of complete data (standalone validation, B2C). B2B engagements (multi-asset, multi-regime) are sized and scheduled at scoping, per the scope and complexity of the strategies validated.

Sample certificate

See Figure 1, 2, or 3 (§5). Each certificate includes the Robustness Score with confidence interval, a verdict with public result band, breakdown by dimension (Edge Quality, Risk Profile, Sample & Coverage), triggered gates with textual justification, detailed metrics supporting each validator, and a cryptographic reproducibility hash for procedural traceability after issuance.

Economic context

TRAVIDENCE pricing is calibrated against the order of magnitude of a challenge fee and the expected cost of failing it: detecting an overfit strategy before you pay turns a negative expected value into an avoided cost. Current pricing at travidence.com.

§7. How to engage

The engagement process has four steps.

1. Upload your data. Accepted formats: TradingView Strategy Tester CSV + Pine Script, MT4/MT5 statement .html, NinjaTrader Trade Performance CSV, Tradovate API (read-only access), ThinkOrSwim trade history, cTrader CSV, or generic broker CSV with timestamps. The evidentiary quality of the input determines the verdict's tier — data exported directly from a platform or broker feed qualifies for T2 (Audit Verified); hand-prepared Excel qualifies for T3 (Conditional Audit).

2. Receive your validation report. Timeline (see §6 for indicative turnaround). The report includes the full Robustness Score with confidence interval, the per-dimension breakdown, the outputs of each of the five validators with their detailed metrics, and the textual justifications for any triggered gates.

3. Get your TRAVIDENCE certificate. PDF cryptographically signed for independent verification, with an sha256 reproducibility hash that lets any third party verify the certificate corresponds exactly to the run over the submitted data with the validators in their declared version.

4. Refine and resubmit if needed. If the verdict is Conditional Robustness or Limited Robustness, the report identifies which dimensions require attention. If it is Not Robust, the report distinguishes between continuous failures (improvable with strategy refinement) and gate failures (typically requiring redesign).

Bilingual support ES + EN throughout the engagement process.

Audit your strategy before you risk capital: travidence.com

§8. Want the full methodology?

This Executive Brief is the retail-facing companion to the D1 Technical Methodology Paper, which documents five novel contributions to the López de Prado validation stack, the full mathematical derivations, the three case studies in institutional detail, the public result band inspired by the procedural discipline of AICPA SAS-122 / ISA 705, and the feature-by-feature comparison against the consumer offering and the implementation libraries. The technical paper is recommended reading if your background includes experience with Combinatorial Purged Cross-Validation (CPCV), the Probability of Backtest Overfitting (PBO), or the López de Prado stack, or if you plan to present TRAVIDENCE pre-validation to institutional counterparties (tier-2 prop firms, family offices, B2B evaluators).

- Arias, Ignacio. Audit-Grade Pre-Validation of Trading Strategies: Five Contributions to the López de Prado Stack (August 17, 2026). Available at SSRN: <https://ssrn.com/abstract=7308022> or <https://doi.org/10.2139/ssrn.7308022>

D1 is recommended reading for institutional quants, B2B counterparties, and reviewers with SSRN experience. This D2 is sufficient reading to make the engagement decision.

References

American Institute of Certified Public Accountants (AICPA). (2011). Statement on Auditing Standards No. 122 — *Statements on Auditing Standards: Clarification and Recodification*. AU-C Section 705: Modifications to the Opinion in the Independent Auditor's Report. New York: AICPA.

International Auditing and Assurance Standards Board (IAASB). (2015). ISA 700 — Forming an Opinion and Reporting on Financial Statements; ISA 705 — Modifications to the Opinion in the Independent Auditor's Report; ISA 706 — Emphasis of Matter Paragraphs and Other Matter Paragraphs in the Independent Auditor's Report. IFAC.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author used Claude (Anthropic) to assist with the drafting and language of the manuscript, under the author's direction. The methodology, analysis, design decisions, and conclusions are the author's own. After using this tool, the author reviewed and edited the content and takes full responsibility for the content of the publication.

Disclaimer

TRAVIDENCE is an independent validation service. It is not financial advice or an investment recommendation. Past performance does not guarantee future results; no validation eliminates the risk of loss.

Ignacio Arias · Travidence, LLC · support@travidence.com Version 1.0 — August 2026 · A TRAVIDENCE public document