MethodsX
Open access
Article Template

Structure-Preserving Randomization for Testing Placement Effects on Exogenous Paths

Aryan Patel

Independent Researcher, United States

Corresponding author: ampatel355@gmail.com | ORCID: 0009-0008-2097-6170

Article information

Article title	Structure-Preserving Randomization for Testing Placement Effects on Exogenous Paths
Authors	Aryan Patel
Affiliations	Independent Researcher, United States
Corresponding author	ampatel355@gmail.com
Keywords	conditional randomization inference; Monte Carlo test; placement effects; permutation test; reproducibility; sensitivity analysis
Related research article	None. The method is presented as a standalone reusable method article.

Highlights

- A structure-preserving randomization method tests whether a realized placement is unusually favorable on a fixed exogenous path.
- The sampler preserves ordered durations, the internal-gap multiset, weights, directions, and non-overlap while re-randomizing placement.
- Validation studies show nominal size under no-skill worlds and monotone power as a placement edge is injected.

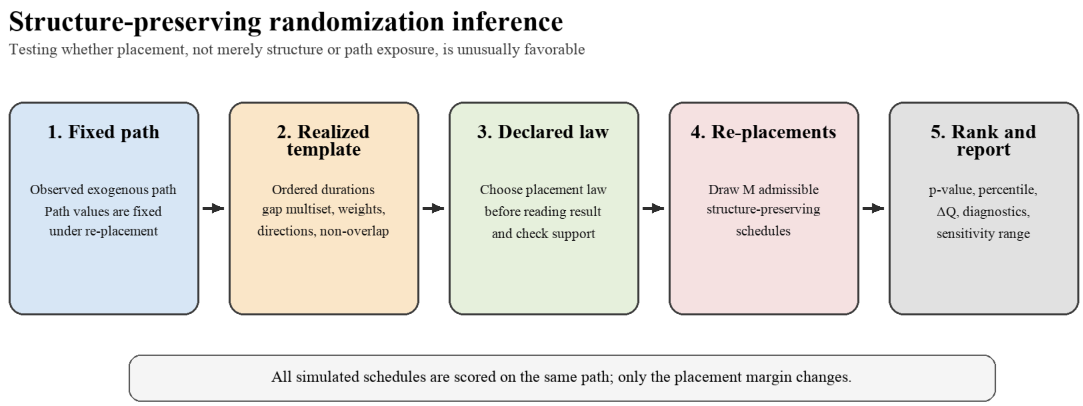
Abstract

Many studies score a fixed sequence of decisions on a realized path that the decisions are assumed not to change. The proposed method asks a narrower question: whether the realized placement of that sequence is unusually favorable after the path and the decision structure are held fixed. The method extracts a structural profile, declares a placement law, draws structure-preserving re-placements, recomputes the statistic on the same path, and reports a plus-one Monte Carlo p-value. The procedure is finite-sample valid under the declared exchangeability law and does not require a model for the path. Validation on synthetic worlds, a small exact-enumeration case, a non-finance monitoring example, and a trading-strategy application illustrates calibration, power, diagnostics, and limits. The method is intended for reusable placement-effect testing rather than as a general test of profitability or total decision quality.

- Compares a realized schedule with structure-preserving re-placements.
- Reports p-value, null percentile, effect size, diagnostics, and sensitivity range.
- Provides implementation checks for support, degeneracy, small samples, and multiplicity.

MethodsX
Open access
Article Template
ELSEVIER

Graphical abstract



Graphical abstract. The method fixes the exogenous path and realized decision structure, re-randomizes placement under a declared law, and reports rank-based evidence with diagnostics.

Background

A realized statistic can look impressive because of the path, because of the structure of the decisions, or because of where that structure landed on the path. These are not the same object. When they are mixed together, a profitable or high-scoring sequence may be read as well timed even when structurally similar placements would have produced the same result.

Structure-preserving randomization separates this placement question from broader performance claims. The exogenous path and the realized decision structure are held fixed. Only the placement of that structure is re-randomized. In a finance application, the path is a price series and the structure is a trade log. In a monitoring application, the path can be a condition process and the structure can be a set of inspection windows. In both cases the inferential question is deliberately narrow: conditional on this structure and this path, does the realized placement appear unusually favorable?

The method is conditional rather than universal. The analyst declares a placement law before reading the result, checks whether the realized schedule lies in its support, draws structure-preserving alternatives, and ranks the realized statistic against those alternatives. Validity follows from the exchangeability of the realized placement with the simulated placements under the declared law, not from a parametric model for the path.

This article provides an implementation-ready method article: required inputs, sampler details, p-value construction, diagnostics, sensitivity analysis, validation studies, and reproducibility guidance. Trading is used as an application, not as the definition of the method.

Method details

Overview of the method

The method has four moving parts. First, a structural profile is extracted from a realized decision sequence. Second, a placement law is declared over feasible re-placements of that same structure. Third, each re-placement is scored on the unchanged exogenous path. Fourth, the realized score is ranked against the simulated scores using a plus-one Monte Carlo p-value.

MethodsX
Open access
Article Template

The procedure is intentionally narrower than a profitability test, model-selection test, or general treatment-quality test. It asks whether the observed placement of a fixed decision template is unusually favorable relative to the placements that the declared law treats as comparable. This narrower scope is what makes the test reusable across domains: trading decisions, inspection windows, monitoring intervals, or other event schedules can be evaluated with the same conditional randomization logic when the path is exogenous to placement.

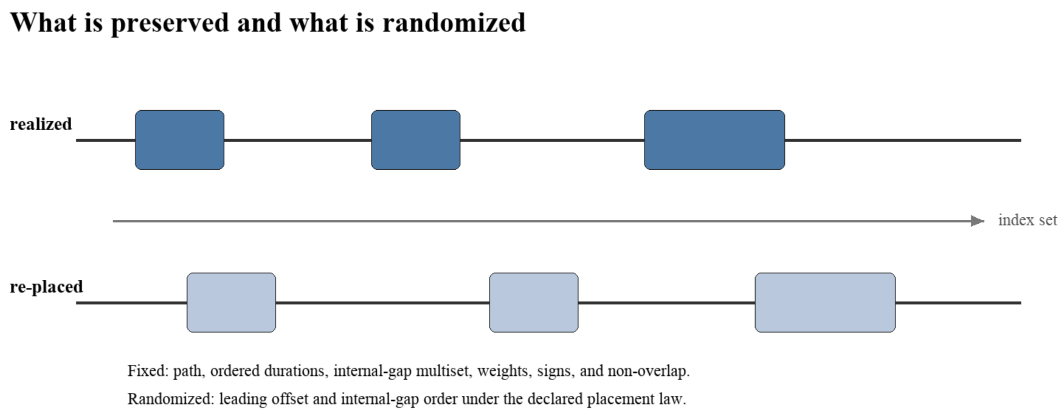


Figure 1. Structure-preserving re-placement. The realized and simulated schedules share the same structural profile and the same path; only the placement margin changes.

Specifications table

Subject area	Mathematics and Statistics; decision support system
More specific subject area	Conditional randomization inference, Monte Carlo testing, computational econometrics, decision-sequence evaluation
Name of the method	Structure-preserving randomization inference for placement effects on exogenous paths
Name and reference of original method	Not a customization of one earlier method. The procedure builds from classical randomization inference [1,2], Monte Carlo significance testing [3-7], and conditional randomization ideas [8].
Resource availability	Code, synthetic-world generators, derived outputs, and figure-generation scripts are available at https://github.com/ampatel355/FORTUNAFRAMEWORK and archived at https://doi.org/10.5281/zenodo.20724999. Raw vendor market data are not redistributed.

MethodsX
Open access
Article Template

Required inputs and outputs

Object	Minimum information	Implementation check
Exogenous path	Ordered values P_t for t = 0,…,T, or any path used by the scoring function	Verify that moving the schedule does not change the path.
Decision log	Entry index, exit index, weight, sign or action label for each decision	Sort by entry time; enforce entry before exit; flag overlaps.
Structural profile	Ordered durations, internal-gap multiset, external slack, weights, signs	Confirm that every simulated schedule preserves these exactly.
Statistic T	Scalar function recomputed on every schedule	Define the statistic before simulation and do not tune it after seeing the realized rank.
Placement law Q_θ	Distribution over feasible schedules preserving the declared structure	Check that the realized schedule lies in support.
Monte Carlo budget M	Number of simulated schedules	Report M and the attainable p-value grid 1/(M + 1).
Outputs	$\hat{p}$, percentile, ΔQ, diagnostics, sensitivity range	Report diagnostics with non-rejections as carefully as with rejections.

Notation and structural profile

Let P_t, t = 0,…,T, be the fixed exogenous path. A realized decision sequence of length N is represented as a set of entry, exit, exposure, and direction records:

$$\mathcal{T} = \{(i^{in}_j, i^{out}_j, \omega_j, d_j)\}_{j=1}^{N}.$$

Here the entry and exit indices are i with superscripts in and out; ω_j is the weight or exposure attached to decision j, and d_j is the sign or action direction. The holding duration and internal gaps are

$$h_j = i^{out}_j - i^{in}_j, \quad h_j > 0,$$

$$g^{int}_j = i^{in}_{j+1} - i^{out}_j, \quad j = 1,\dots,N - 1.$$

External slack is the unused room before the first decision and after the last decision:

$$g^{ext} = i^{in}_1 + (T - i^{out}_N).$$

The structural profile is therefore

$$\mathcal{S} = (h, g^{int}, g^{ext}, \omega, d).$$

The method preserves this profile. Only the calendar placement changes.

Symbol	Meaning
P□	Exogenous path value at index t
N	Number of decisions in the realized sequence
h□	Holding duration of decision j
g^{int}□	Internal gap between decisions j and j+1
g^{ext}	Total leading and trailing slack
ω□, d□	Weight and sign/direction attached to decision j
S_0	Realized schedule
S□	mth structure-preserving re-placement
T(S)	Statistic computed from schedule S on the fixed path
Q_θ	Declared placement law
$\hat{p}$	Plus-one Monte Carlo p-value using M re-placements
ΔQ	Realized statistic minus the simulated mean under Q_θ

Assumptions

- Path exogeneity: moving the decision schedule must not change the realized path.
- Exchangeability: under the no-placement-skill null, the realized placement is treated as one draw from the same conditional law used to generate simulated placements.
- Support: the realized schedule must be drawable under the declared placement law.

Algorithm 1. Core structure-preserving randomization workflow

Input: Exogenous path P□ for t = 0,…,T; realized decision log; statistic T(.); placement law Q_θ; Monte Carlo budget M.

Output: Plus-one p-value $\hat{p}$, null percentile, effect estimate ΔQ, and diagnostics.

1. Clean and align the path and decision log.
2. Extract $\mathscr{S} = (h, g^{int}, g^{ext}, \omega, d)$ from the realized log.
3. Check non-overlap, positive durations, and realized-schedule support under Q_θ.
4. Compute the observed statistic $T(S_0)$ on the fixed path.
5. For m = 1,…,M, draw S□ from Q_θ while preserving $\mathscr{S}$ and compute T□ = T(S□).
6. Set k = Σ□ 1{T□ ≥ $T(S_0)$}.
7. Report $\hat{p} = (1 + k)/(M + 1)$, percentile 1 - $\hat{p}$, and ΔQ = $T(S_0)$ - mean□ T□.
8. Run diagnostics and, if planned, repeat the calculation over the declared placement-law menu.

Algorithm 2. Gap-permutation sampler Q_{gp}

Input: Ordered durations h□ for j = 1,…,N; internal gaps g^{int}□ for j = 1,…,N - 1; and external slack g^{ext}.

Output: A feasible schedule S preserving the realized structural profile.

1. Draw a leading gap uniformly from $\{0,\dots,g^{ext}\}$.
2. Draw a uniformly random permutation of the realized internal-gap multiset.
3. Set the first entry after the drawn leading gap and preserve h_1.
4. Place each later decision after the previous exit plus the next permuted internal gap.
5. Preserve every ordered duration, weight, sign, and the non-overlap constraint.

The feasibility reason is short. The total internal span, equal to the sum of durations plus the sum of internal gaps, does not depend on the gap permutation. Since external slack is the total unused leading and trailing room, every leading gap from 0 to g^{ext} leaves nonnegative trailing slack.

$$\text{total span} = \Sigma^{N}_{j=1} h_j + \Sigma^{N-1}_{j=1} g^{int}_j.$$

Statistic, p-value, and effect summaries

The statistic T(S) must be selected before running the simulation. In the trading example T(S) is cumulative return, but the procedure only requires a scalar score that can be recomputed on each feasible schedule.

The one-sided p-value is

$$\hat{p}_M = (1 + \Sigma^M_{m=1} 1\{T(S_m) \geq T(S_0)\}) / (M + 1).$$

The plus-one correction keeps the Monte Carlo test valid at finite M and prevents an impossible zero p-value [6]. The matching effect estimate is

$$\Delta Q = T(S_0) - (1/M)\, \Sigma^M_{m=1} T(S_m).$$

The null percentile and null dispersion should also be reported. A standardized score using the simulated mean μ and standard deviation σ is useful descriptively, but it should not be treated as a normal z-test unless additional central-limit conditions are justified.

Sensitivity analysis over placement laws

Measure	Preserved features	Question answered
Gap permutation	Path, ordered durations, gap multiset, weights, signs, non-overlap	Was the realized placement unusually favorable conditional on the realized spacing burden?
Leading-gap restricted	Same as gap permutation, plus a common warm-up exclusion	Was placement favorable after removing calendar regions unavailable to the rule?
Context- or regime-matched	Same as gap permutation, plus entry-time state features available at entry	Was placement favorable beyond matching the local state?
Uniform feasible schedules	Path, durations, weights, signs, non-overlap; gap multiset relaxed	A looser spacing null, useful as a contrast rather than the headline law.
Slack redistribution	Path, durations, weights, signs, approximate turnover; exact gap structure relaxed	A stress test for how much exact gap preservation matters.

Algorithm 3. Sensitivity range over a declared menu

Input: Declared admissible menu Θ_0; path; structural profile; statistic; Monte Carlo budget M.

Output: Per-law p-values, effect estimates, sensitivity range, and measure-invariant verdict.

1. For each θ in Θ_0, draw M schedules from Q_θ.
2. Compute the law-specific p-value $\hat{p}$ and effect estimate ΔQ.
3. Run support and degeneracy diagnostics for each law.
4. Report the minimum and maximum $\hat{p}$ values across Θ_0 and the corresponding ΔQ range.
5. Declare measure-invariant rejection at level α only if every law-specific $\hat{p} \leq \alpha$.

Diagnostics

Diagnostic	What it detects	Recommended response
Support check	The realized schedule is not drawable under the declared law	Do not report a valid randomization p-value under that law.
Null dispersion	Simulated statistics are nearly identical	Report weak identification; avoid strong claims from non-rejection.
Small-N warning	Few decisions imply little placement information	Prefer exact enumeration where possible; treat borderline ranks cautiously.
Dominant-decision ratio	One decision carries most of the statistic	Report percentile rather than relying on a standardized normal score.
Measure sensitivity	The verdict appears only under one law	Present as law-dependent, not robust placement evidence.
Multiplicity	Many schedules, rules, or assets are tested	Apply a planned correction such as Benjamini-Hochberg or Bonferroni.

Computational considerations

For one strategy, the sampler cost is O(MN) because each of M schedules reassembles and scores N decisions. Memory can be kept at O(M) if only the simulated statistic vector is stored, or O(1) if only the running count and running moments are retained. Storing the vector is useful because it supports histograms, percentiles, and null-dispersion diagnostics.

Exact enumeration is preferable when the orbit is small. Under the gap-permutation law an upper bound is $(g^{ext} + 1)(N - 1)!$, with fewer distinct schedules when internal gaps contain ties. This grows quickly, so enumeration is mainly useful for small N.

Algorithm 4. Validation recipe for a new domain

Input: Domain-specific path generator or historical path collection; decision-template generator; scoring statistic.

Output: Size, power, degeneracy, and small-sample evidence for the method in that domain.

1. Generate no-skill schedules from the same placement law used by the null.
2. Run the test and check that null p-values are close to uniform.
3. Generate positive-control schedules with a known placement edge.
4. Verify that rejection probability rises as the injected edge increases.
5. Run at least one degeneracy case with flat or near-flat paths.
6. Run at least one small-N case and compare Monte Carlo against exact enumeration when feasible.
7. Document software versions, seeds, data exclusions, and all chosen measures.

MethodsX
Open access
Article Template

Method validation

Synthetic calibration and power

The method was validated on synthetic worlds where the truth is known. In no-skill worlds, the realized placement is generated without placement information, so the test should reject at the nominal rate. In positive-control worlds, a placement edge is injected, so rejection should increase with signal strength.

World or signal	Mean actual return	Reject at 5%	KS p vs. uniform	Mean standardized effect
IID, no skill	-0.00371	0.05133	0.5153	-0.00924
Volatility clustering, no skill	0.00337	0.04933	0.7009	0.00884
Drift only, no skill	0.15531	0.05600	0.5978	0.01788
Structural exposure, no timing	0.62207	0.05467	0.5662	0.00724
0 bp signal	0.00164	0.0425	0.9424	0.01027
5 bp signal	0.15833	0.1300	7.65e-34	0.69904
10 bp signal	0.38461	0.4750	2.17e-117	1.73153
20 bp signal	0.92064	0.9250	4.66e-280	3.79852
35 bp signal	2.08394	1.0000	0.0	7.22876

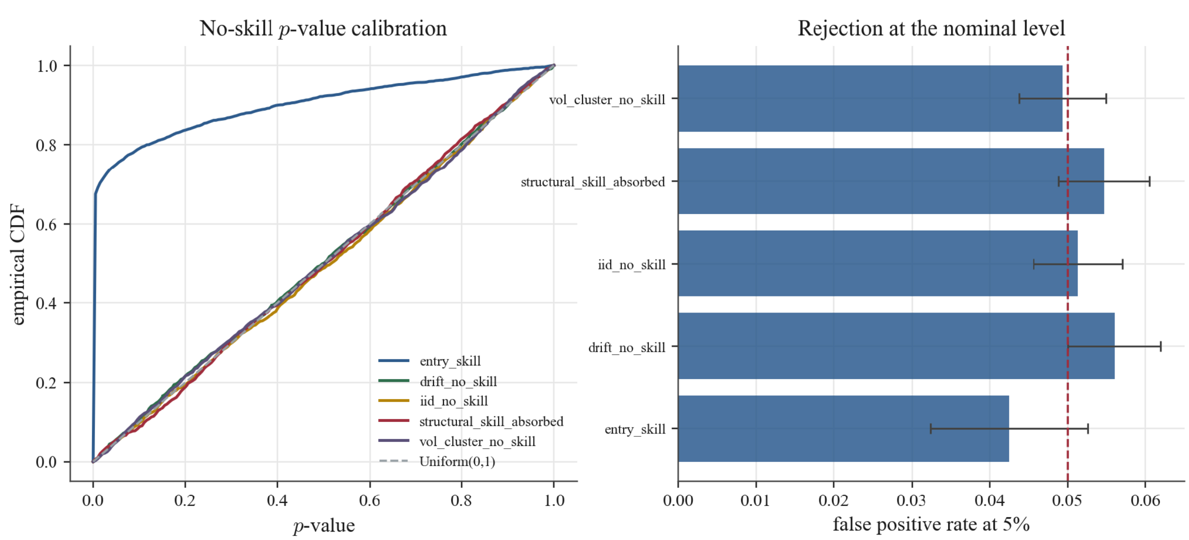


Figure 2. Synthetic no-skill calibration. The null p-value distribution is close to uniform, and rejection rates remain near the nominal 5% level.

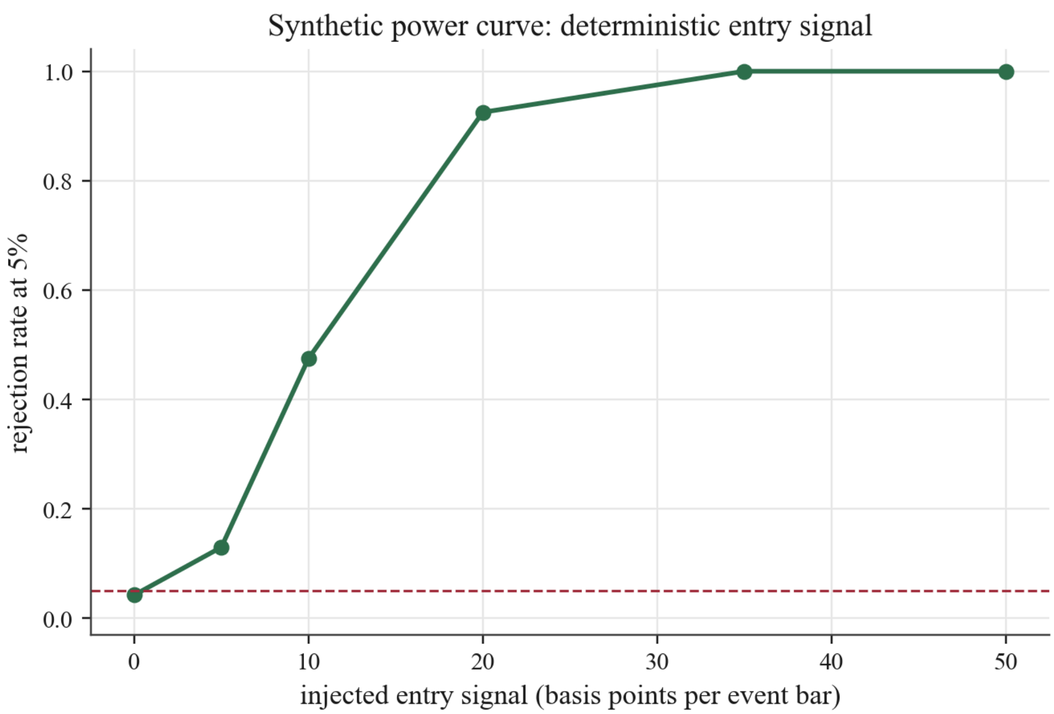


Figure 3. Synthetic power curve. Rejection rises monotonically as a deterministic placement signal is injected.

Comparison with profitability tests

A useful validation case is a strategy that is profitable but not timed. In that world, a profitability test should often reject, while a placement test should not. The controlled comparison with White's Reality Check [9] and Hansen's SPA [10] shows this distinction.

World	Reality Check	SPA	Placement test
IID, no skill	0.138	0.128	0.052
Drift only, no skill	0.388	0.393	0.060
Volatility clustering, no skill	0.125	0.122	0.062
Structural exposure, no timing skill	0.660	0.637	0.050
Genuine entry-placement skill	1.000	1.000	1.000

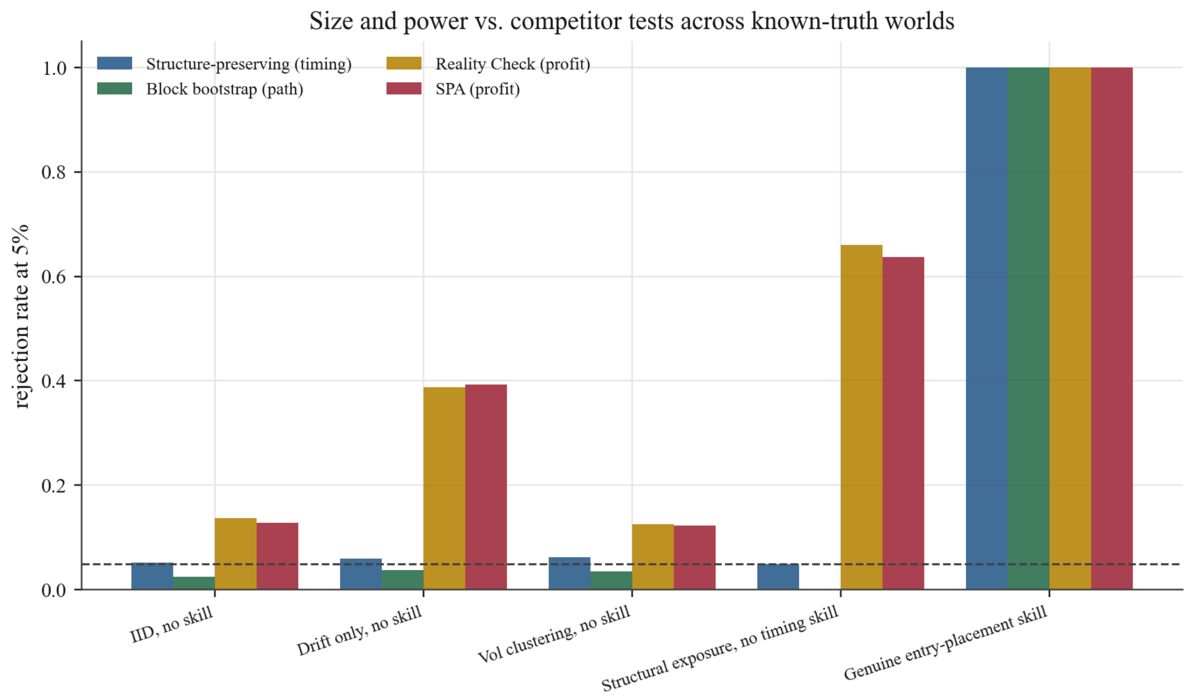


Figure 4. Known-truth comparison. In the structural-exposure world, profitability tests reject while the placement test holds size.

Exact enumeration at small decision counts

Where the orbit is small enough, Monte Carlo can be checked against exact enumeration. For an $N = 4$ schedule with durations [1, 2, 4, 6], internal gaps [834, 4094, 637], and external slack 470, the gap-permutation orbit contains $3! \times (470 + 1) = 2826$ feasible schedules. The exact one-sided tail is $q = 0.057325$, while a same-cost 20,000-draw Monte Carlo estimate gives 0.059747. Both values lead to the same non-rejection.

$$3! \times (470 + 1) = 2826 \text{ feasible schedules.}$$

Non-finance validation

The same sampler was applied to condition-monitoring windows. The exogenous path was a condition process, the structural profile was a set of monitoring-window durations and gaps, and the statistic was the condition change captured inside active windows. Under uninformative placement the test holds size; under forecast-guided placement, power rises as the forecast improves.

Experiment	N	Forecast noise	Rejection at 0.05
Size	4	---	0.0505 (0.0035)
Size	8	---	0.0450 (0.0033)
Size	16	---	0.0428 (0.0032)
Size	32	---	0.0490 (0.0034)
Power	16	0.5	0.9955 (0.0011)
Power	16	1.0	0.8433 (0.0057)
Power	16	2.0	0.4492 (0.0079)
Power	16	4.0	0.2057 (0.0064)

Example application: trading-strategy timing

The finance application is included as an example, not as the definition of the method. On the gold-futures panel, ten of eleven rules earned positive cumulative return, but no rule crossed $p \leq 0.05$ under the neutral gap-permutation measure. On the cross-asset scan, 13 of 322 tests had nominal $p \leq 0.05$, fewer than the 16.1 expected by chance, and no test survived Benjamini-Hochberg or Bonferroni correction.

Table. Gold-futures eleven-rule example under the neutral gap-permutation measure. Cumulative returns are net of round-trip transaction cost c = 0.000470; the random-entry row is a single seed-42 realization.

Rule	Trades	Cum. return	p	Std. effect	Percentile
Trend Pullback	77	0.083	0.870	-1.03	13.0
Breakout Vol+Mom	46	0.297	0.176	0.91	82.4
Mean Rev Vol Filter	13	-0.010	0.662	-0.44	33.8
Validation AVM	35	0.219	0.345	0.32	65.5
ADX Trend	110	0.803	0.141	1.08	85.9
Oversold Reversion	30	0.082	0.341	0.35	65.9
Squeeze Breakout	4	0.065	0.053	1.50	94.7
Connors RSI2	41	0.118	0.143	1.07	85.7
Donchian Reentry	60	0.071	0.647	-0.46	35.3
Turn-of-Month	141	0.323	0.234	0.65	76.7
Random control	215	0.601	0.710	-0.64	29.0
Buy and hold	---	15.944	---	---	---

Quantity	Value
Instruments	47
Asset x rule tests	322
Nominal $p \leq 0.05$ observed / expected by chance	13 / 16.1
Benjamini-Hochberg discoveries	0
Bonferroni discoveries	0
KS statistic vs. uniform (p-value)	0.114 (0.0004)
Smallest panel p-value	0.0014
Random-entry baselines in smallest 10	2

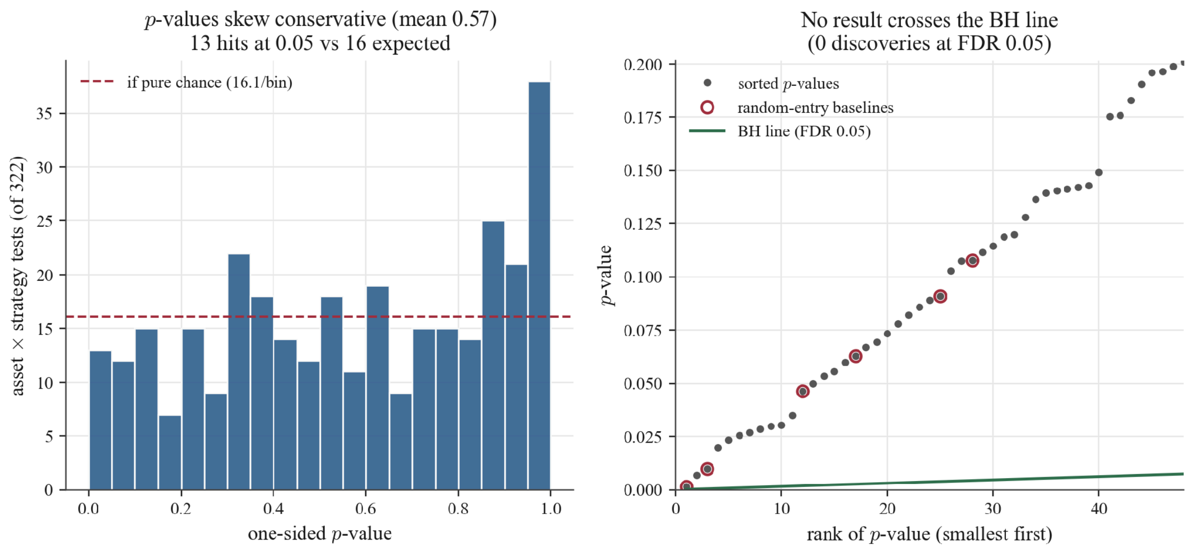


Figure 5. Cross-asset multiplicity scan. The panel has fewer nominal rejections than expected by chance, and no discoveries survive standard multiplicity correction.

Limitations

The method is conditional, and that is both its strength and its limit. It is strong because the comparison is exact under the declared exchangeability law. It is limited because anything held fixed by the structural profile is no longer being tested. If the useful part of a strategy is its holding-period rule, exit logic, sizing rule, or decision to skip events, a placement test that conditions on those objects will not discover that skill.

The path must be exogenous to the moved decisions. The method should not be used for interventions that change the future path unless a separate causal model justifies reusing the same path under all placements.

Small decision counts reduce power. A four-decision schedule can still be tested, and sometimes enumerated exactly, but the test should be reported as a weakly identified placement comparison. A non-rejection in that setting should not be turned into a claim that no skill exists.

Finally, the placement law is analyst-declared. A rejection under one law and not another is a law-dependent result, not a universal discovery. This is why the method reports sensitivity ranges and diagnostics rather than a single polished number.

Ethics statements

This work uses synthetic simulations, publicly accessible software metadata, and derived market-data outputs. It does not involve human participants, animal experiments, or social-media data. No ethics approval or informed consent was required.

CRediT author statement

Aryan Patel: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review and editing, Visualization.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of interests

The author declares no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During preparation of this work, the author used OpenAI Codex and Claude Opus 4.8 to assist with generating code used for simulations and figures, and to review grammar, wording, and clarity. After using these tools, the author reviewed and edited the material as needed and takes full responsibility for the content of the article.

Supplementary material and/or additional information

The GitHub repository contains the Python implementation, synthetic validation scripts, figure-generation scripts, derived data files, and archived outputs needed to reproduce the tables and figures. Raw vendor market data are not redistributed and remain subject to the terms of the original data provider.

Appendix A. Finite-sample validity

Fix a declared admissible placement law Q_θ. Let M independent schedules $S_1, S_2, \ldots$ be drawn from Q_θ conditional on the fixed path and structural profile. Under the sharp no-placement-skill null, the realized schedule S_0 is also a draw from that same law. Therefore the realized statistic and the simulated statistics are exchangeable conditional on the path and structure:

$$T(S_0), T(S_1), \ldots, T(S_M).$$

If there are no ties, the rank of $T(S_0)$ among the $M + 1$ values is uniform on $\{1,\ldots,M + 1\}$. Counting ties into the right tail can only make the test conservative. Hence, for grid values α in $\{1/(M + 1),\ldots,1\}$,

$$\Pr(\hat{p}_M \leq \alpha \mid \mathcal{C}) \leq \alpha.$$

This finite-sample guarantee does not require a model for returns, normality, or a large-sample approximation. It uses exchangeability of the realized placement with the simulated placements. For fixed S_0, large-M consistency follows from the strong law of large numbers because the indicators $1\{T(S_{\square}) \geq T(S_0)\}$ are independent Bernoulli draws conditional on S_0 and the conditioning information.

Appendix B. Implementation checklist

1. Define the exogenous path and justify why the path would not change under re-placement.
2. Define the realized decision log and remove impossible records before simulation.
3. Extract the ordered durations, internal-gap multiset, external slack, weights, and signs.
4. Declare the statistic before seeing the simulated null.
5. Declare the placement law and confirm that the realized schedule lies in its support.
6. Set M and record the random seed.
7. Generate schedules, score them on the unchanged path, and compute the plus-one p-value.

8. Report the percentile, effect estimate, null dispersion, and all warnings.
9. Repeat over the planned measure menu if sensitivity analysis is part of the claim.
10. Archive code, derived outputs, and environment information.

References

[1] R. A. Fisher, The Design of Experiments, Oliver & Boyd, Edinburgh, 1935.
[2] E. L. Lehmann, J. P. Romano, Testing Statistical Hypotheses, 3rd ed., Springer, New York, 2005.
[3] M. Dwass, Modified randomization tests for nonparametric hypotheses, Annals of Mathematical Statistics 28 (1957) 181-187.
[4] A. C. A. Hope, A simplified Monte Carlo significance test procedure, Journal of the Royal Statistical Society: Series B 30 (1968) 582-598.
[5] J. Besag, P. Clifford, Generalized Monte Carlo significance tests, Biometrika 76 (1989) 633-642.
[6] B. Phipson, G. K. Smyth, Permutation p-values should never be zero, Statistical Applications in Genetics and Molecular Biology 9 (2010) Article 39.
[7] J. Hemerik, J. Goeman, Exact testing with random permutations, TEST 27 (2018) 811-825.
[8] E. Candes, Y. Fan, L. Janson, J. Lv, Panning for gold: model-X knockoffs for high-dimensional controlled variable selection, Journal of the Royal Statistical Society: Series B 80 (2018) 551-577.
[9] H. White, A reality check for data snooping, Econometrica 68 (2000) 1097-1126.
[10] P. R. Hansen, A test for superior predictive ability, Journal of Business and Economic Statistics 23 (2005) 365-380.
[11] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate, Journal of the Royal Statistical Society: Series B 57 (1995) 289-300.