

Edge-Persistenz: Überlebt ein gemessener Edge — und welcher Teil davon?

10 Märkte · 2015–2026

123.193 Trades über 10 Märkte und 11,4 Jahre, zerlegt in strukturellen Edge, Marktregime und frische Auffälligkeit. Der strukturelle Edge trägt über fünf Jahre ($\beta=0,90$). Eine frische Auffälligkeit verliert 89 % ihrer Größe in einem einzigen Halbjahr. Die Vergangenheit ist ein Forecast — nur nicht der Teil, auf den alle schauen.

Dokument-Metadaten

Feld	Wert
Herausgeber	KruegerAlgorithms
Titel	Edge-Persistenz: Überlebt ein gemessener Edge — und welcher Teil davon?
Erscheinungsdatum	27. Juli 2026
Version	v1.0
Verdikt	Edge bestätigt
Märkte	10 Märkte
Zeitraum	2015–2026
Stichprobe	123.193 Episoden
Datengrundlage	Datenbasis: 123.193 Episoden, 10 Märkte (DAX, FTSE, DOW, NQ, SPX, HK50, JPN225, XAUUSD, EURUSD, GBPUSD), 23 Setups, 05.01.2015 – 05.06.2026 (11,4 Jahre, 23 Halbjahre). Exit: Trailing-Stop BE 0,5 / TS 1,0 / Schritt 0,5, netto nach Spread und Slippage . Einheit: R (Risikovielfaches). Keine Handelsempfehlung.
Dokumenttyp	Statistische Auswertung (keine Trading-Empfehlung)

Rein statistische Auswertung historischer Marktdaten. Keine Trading-Empfehlung, keine Finanzberatung. Historische Performance ist keine Garantie für die Zukunft. Trading mit Hebelprodukten kann zum Totalverlust führen.

Zusammenfassung

123.193 Trades über 10 Märkte und 11,4 Jahre, zerlegt in strukturellen Edge, Marktregime und frische Auffälligkeit. Der strukturelle Edge trägt über fünf Jahre ($\beta=0,90$). Eine frische Auffälligkeit verliert 89 % ihrer Größe in einem einzigen Halbjahr. Die Vergangenheit ist ein Forecast — nur nicht der Teil, auf den alle schauen.

Jeder Trader kennt die Situation. Ein Setup, das ein halbes Jahr lang sauber lief, hört auf zu funktionieren. Ein Markt, der tot war, wacht auf. Dahinter steht immer dieselbe Frage: **Wenn ich einen Edge messe — was sagt mir das über die Zukunft, und sagt es mir überhaupt etwas?**

Die meisten Antworten darauf sind Anekdoten. Das hier ist eine Messung. Sie fragt, ob die Vergangenheit ein Forecast ist, und die Antwort lautet ja — aber für einen anderen Teil der Daten als den, auf den die meisten schauen.

Diese Auswertung enthält keine Trading-Empfehlung.

Inhaltsverzeichnis

Zusammenfassung	2
Inhaltsverzeichnis	3
Abbildungen	3
1. Das Modell: in einer Zahl stecken drei Dinge	4
2. Wie viel an einem gemessenen Edge ist überhaupt echt?	4
3. Die Auffälligkeit zerfällt nicht — sie stürzt von der Klippe	5
4. Der strukturelle Edge dagegen hält	5
5. Das Alter der Information ist fast egal	6
6. Zwei Prüfungen, die unsere eigenen Hypothesen erledigt haben	7
7. Was daraus für die Praxis folgt	8
8. Grenzen	9
Die Ein-Satz-Fassung	9

Abbildungen

Abb. 1 — variance-decomposition.png

Abb. 2 — decay-curve.png

Abb. 3 — split-half.png

Abb. 4 — information-age.png

Abb. 5 — selection-rules.png

1. Das Modell: in einer Zahl stecken drei Dinge

Wenn eine Zelle — ein Setup, in einem Markt, in einer Richtung — in einem Halbjahr +0,30 R abwirft, enthält diese Zahl vier Dinge gleichzeitig:

gemessen = struktureller Edge + Marktregime + frische Auffälligkeit + Stichprobenrauschen

- **Struktureller Edge** (α): was diese Zelle *dauerhaft* wert ist. Über die Zeit konstant.
- **Marktregime** (τ): die gemeinsame Welle, die in dieser Periode alle Zellen hebt oder senkt.
- **Frise Auffälligkeit** (ε): diese Zelle verhält sich *in genau dieser Periode* anders als sonst. Das ist gemeint, wenn jemand sagt „das Setup läuft gerade heiß“.
- **Stichprobenrauschen**: reiner Zufall aus einer endlichen Zahl von Trades.

Mit dem Auge sind die vier nicht trennbar, rechnerisch schon — über zweifaches Mittelwert-Abziehen entlang beider Achsen, 85 Zellen $\times$ 23 Halbjahre. Und sie verhalten sich **vollkommen unterschiedlich**. Genau dieser Unterschied ist der Kern der Studie.

2. Wie viel an einem gemessenen Edge ist überhaupt echt?

In der beobachteten Streuung steckt immer Stichprobenrauschen. Dessen Größe ist aber bekannt — jede Zell-Periode hat ihren eigenen Standardfehler. Zieht man den mittleren quadrierten Standardfehler von der beobachteten Varianz ab, bleibt das echte Signal übrig.

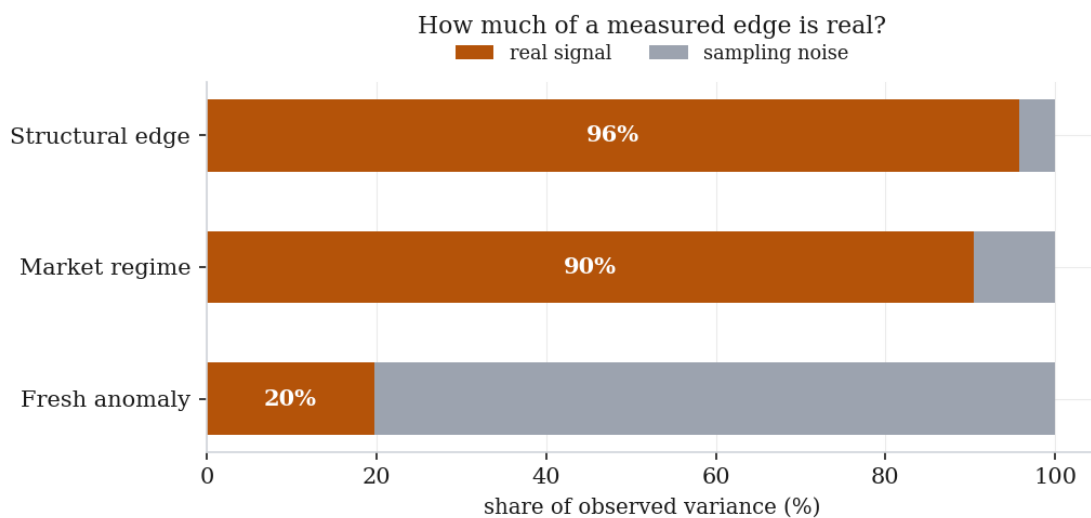


Abb. 1:

Komponente	SD beobachtet	SD Rauschen	SD echt	Echtes Signal
Struktureller Edge	0,162	0,033	0,159	96 %
Marktregime	0,055	0,017	0,052	90 %
Frise Auffälligkeit	0,174	0,156	0,077	20 %

Die letzte Zeile lohnt genaues Lesen. Wenn eine Zelle in einem Halbjahr von ihrer eigenen Norm abweicht, sind **80 % dieser Abweichung Rauschen**. Die Unterschiede zwischen Setups sind fast vollständig echt. Die Unterschiede zwischen *Perioden innerhalb desselben Setups* sind es fast vollständig nicht.

Bemerkenswert: Die beobachtete Streuung ist in Zeile eins und drei praktisch gleich groß (0,162 vs. 0,174). Auf dem Bildschirm sehen ein struktureller Edge und eine frische Auffälligkeit **exakt gleich groß aus**. Nur die Zerlegung unterscheidet sie.

3. Die Auffälligkeit zerfällt nicht — sie stürzt von der Klippe

Man nehme jede Zelle, die in einer Periode auffällig wird ($|t| \geq 2$ gegen ihre eigene Norm), und verfolge sie weiter.

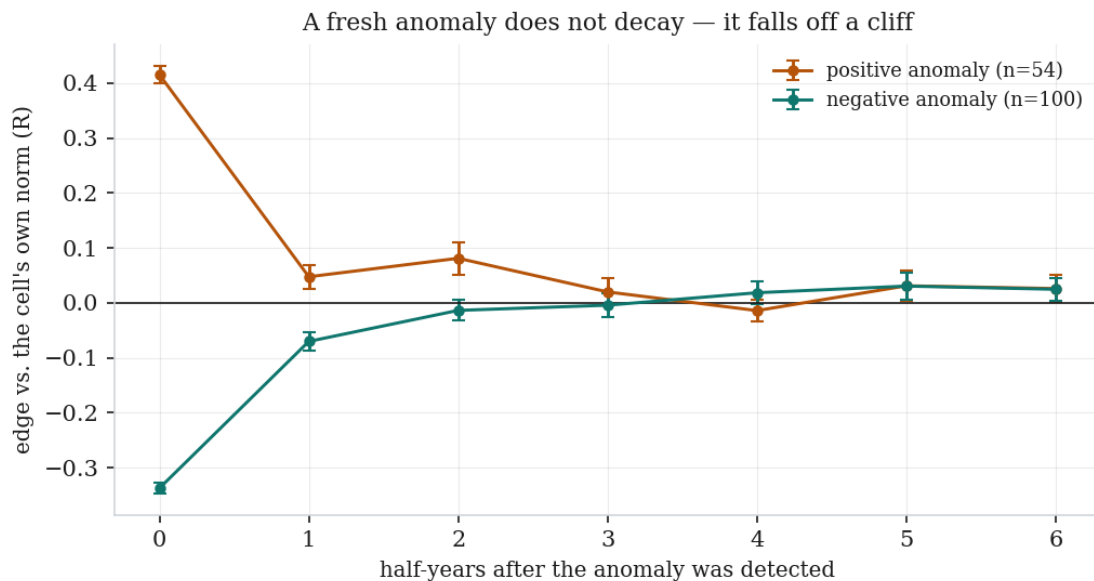


Abb. 2:

Halbjahre danach	0	1	2	3	4	5	6
nach positiver Auffälligkeit (n=54)	+0,416	+0,048	+0,081	+0,020	-0,014	+0,031	+0,026
nach negativer Auffälligkeit (n=100)	-0,337	-0,070	-0,014	-0,004	+0,019	+0,031	+0,025

89 % des Effekts sind nach einem Halbjahr weg. Nicht allmählich — sofort. Und es ist symmetrisch: Eine Zelle, die gerade fürchterlich lief, ist genauso schnell wieder auf ihrem Normalniveau.

Das ist in beide Richtungen relevant, und die zweite ist die teure. Der Reflex, eine Strategie nach einer schlechten Strecke abzuschalten, beruht auf demselben Denkfehler wie der Reflex, auf eine heiße aufzuspringen. Keine der beiden Strecken sagt viel über die nächste.

Die Form widerlegt außerdem das intuitive Modell. Man erwartet, dass ein Edge sich „abnutzt“ — entdeckt, wegarbitriert, allmählich zerfallend. Was wir sehen, ist keine Zerfallskurve. Es ist eine Stufe: ein großer gemessener Wert, dann ein kleiner, flacher, ungefähr konstanter Rest. Das ist der Fingerabdruck eines **Selektionseffekts**, nicht der von Erosion.

4. Der strukturelle Edge dagegen hält

Dieselben Daten, andere Frage. Die gesamte Historie in zwei Hälften teilen, je Zelle den Mittelwert in beiden Hälften berechnen und gegeneinander auftragen.

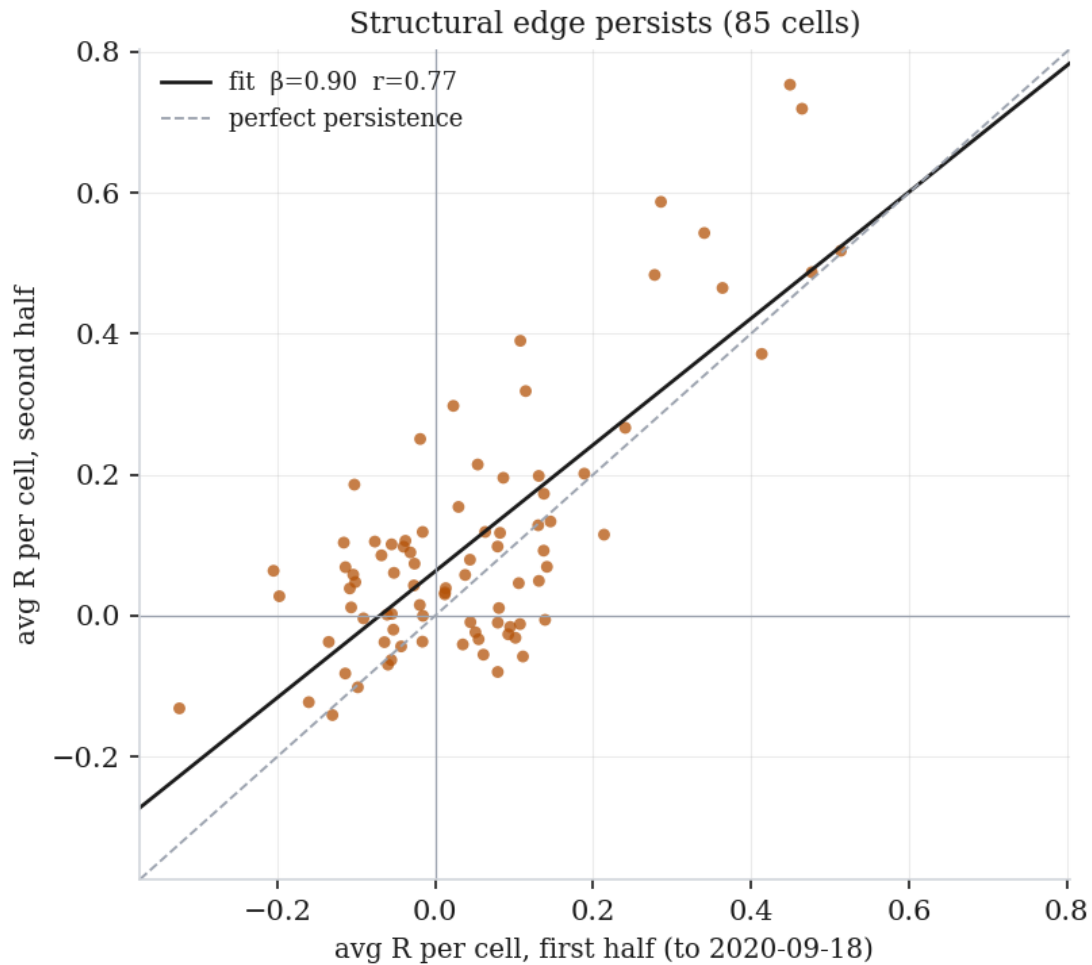


Abb. 3:

$\beta = 0,90$ (SE 0,082, $t = 10,9$), $r = 0,77$, 85 Zellen. Eine Steigung von 0,90 heißt: Eine Zelle, die in der ersten Hälfte 0,10 R besser war als der Durchschnitt, war es in der zweiten immer noch um rund 0,09 R — über einen Abstand von fast fünf Jahren.

Erste Hälfte (bis 18.09.2020)	→	Zweite Hälfte
Top-Quintil +0,288	→	+0,344
Alle Zellen +0,046	→	+0,110
Bottom-Quintil -0,129	→	+0,003

Das Top-Quintil bleibt oben. Das Bottom-Quintil bleibt *nicht* unten — es kehrt auf etwa null zurück, was für sich genommen aufschlussreich ist: Dauerhaft *schlechte* Zellen sind seltener als dauerhaft gute.

5. Das Alter der Information ist fast egal

Würde ein Edge über die Zeit zerfallen, müssten frische Daten besser prognostizieren als alte. Sie tun es nicht.

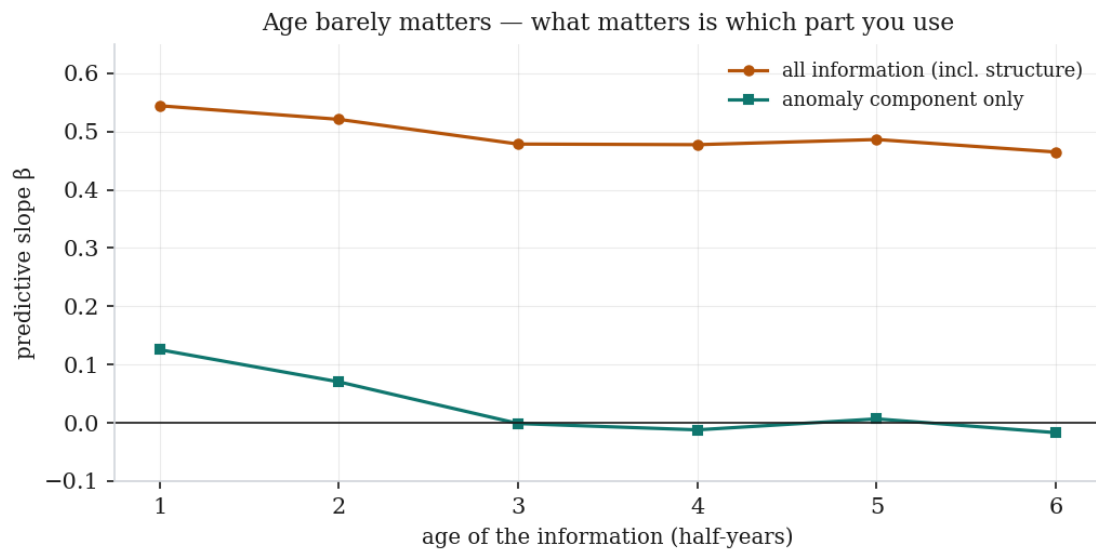


Abb. 4:

Alter (Halbjahre)	1	2	3	4	5	6
gesamte Information	0,544	0,521	0,479	0,478	0,486	0,465
nur Auffälligkeits-Anteil	0,126	0,070	-0,001	-0,012	+0,007	-0,017

Ein Halbjahr von **vor drei Jahren** prognostiziert die nächste Periode fast so gut (0,465) wie das gerade abgelaufene (0,544). Der Informationsgehalt ist im Alter praktisch flach — weil er von der strukturellen Komponente getragen wird, und der ist egal, wie alt sie ist.

Der Abstand zwischen den beiden Zeilen ist das ehrliche Maß für den Wert von Aktualität: $0,544 - 0,465 \approx 0,08$. So viel fügt „was gerade passiert ist“ zu „was immer schon galt“ hinzu.

Der Auffälligkeits-Anteil ist bei Lag 1 nicht null (+0,126). Gegen eine Permutations-Nullverteilung geprüft — Perioden innerhalb jeder Zelle gemischt, 500 Durchläufe — liegt er bei $z = 6,6$. Es gibt also ein echtes, winziges Momentum in frischen Auffälligkeiten. Es ist statistisch belastbar und praktisch fast bedeutungslos, und ab Lag 3 ist es verschwunden.

6. Zwei Prüfungen, die unsere eigenen Hypothesen erledigt haben

Regime-Momentum überlebt die Trendbereinigung nicht. Auf gute Halbjahre scheinen gute Halbjahre zu folgen: AR(1) auf dem Regime-Faktor ergibt $\beta = +0,71$ ($t = 3,9$). Der Regime-Faktor hat über den Zeitraum aber einen leichten Aufwärtstrend, und ein Trend allein erzeugt bereits positive Autokorrelation. Trendbereinigt bricht es auf $\beta = +0,29$ ($t = 1,4$) ein — nicht signifikant. Was wie „Momentum in den Marktbedingungen“ aussah, war überwiegend ein Trendartefakt.

Edges erodieren nicht. Die naheliegende Sorge: Werden Märkte effizienter, müsste der Abstand zwischen guten und schlechten Zellen schrumpfen.

Epoche	avg R	SD echt (Streuung zwischen Zellen)
2015–2018	+0,024	0,145
2019–2022	+0,100	0,189
2023–2026	+0,115	0,164

Keine Erosion über 11 Jahre — eher das Gegenteil. **Vorbehalt, und ein wichtiger:** Unsere Setup-Familie wurde mit Kenntnis der jüngeren Jahre entwickelt, die späteren Epochen sind also teilweise In-Sample. Die ehrliche

Lesart ist „kein Hinweis auf Erosion“, nicht „Beweis, dass es keine gibt“.

Was diese Studie uns gekostet hat. Die Analyse hat einen Nebeneffekt zutage gefördert, nach dem wir nicht gesucht hatten: einen scheinbaren „Ähnliche-Tage“-Effekt von +0,73 R, der zu gut war. Es war ein **Lookahead-Leck** — zwei Kontext-Features (Opening-Range-Level und ein aus dem Cash-Open abgeleiteter Bias) wurden gelesen, bevor sie bekannt sein konnten, in 48,8 % der Episoden. Nach dem Fix fiel der Effekt auf +0,19 R. Die Diagnose, die es entlarvt hat: Die Prognose *unter*trieb das Ergebnis — und ein verrauschter Prädiktor muss immer *über*treiben. Alle R-Werte dieses Papers sind unberührt: Der Bug betraf Features, nie Ergebnisse, und nach dem Neubau reproduzierten alle Anker bitgenau. Wir erwähnen es, weil ein Paper über Messfehler, das seinen eigenen verschweigt, wenig wert wäre.

7. Was daraus für die Praxis folgt

Vier Auswahlregeln, walk-forward: Zu jeder Periode werden Zellen nur mit Daten ausgewählt, die *davor* verfügbar waren, und dann für diese Periode gehalten.

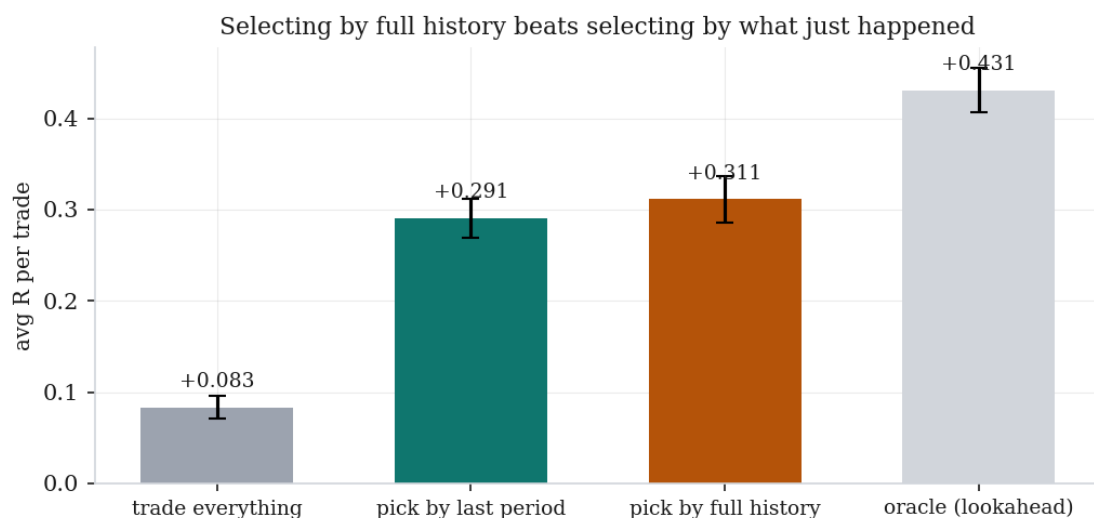


Abb. 5:

Regel	avg R pro Trade
Alles handeln	+0,083
Top-Quintil nach letzter Periode	+0,291
Top-Quintil nach ganzer Historie	+0,311
Orakel (perfekte Voraussicht)	+0,431

Zwei Dinge fallen auf. Erstens funktioniert die Auswahl nach Aktualität überraschend gut (+0,291) — aber nicht, weil Aktualität informativ wäre. Sie funktioniert, weil die Zahl der letzten Periode ein *verrauschter Stellvertreter* für den strukturellen Edge ist. Die ganze Historie direkt zu nutzen erledigt dieselbe Aufgabe besser.

Zweitens beträgt der Abstand zwischen der besten nutzbaren Regel (+0,311) und perfekter Voraussicht (+0,431) nur 0,12 R. **Das ist der gesamte Raum, in dem jede adaptive, timing-basierte Intelligenz operieren kann** — und wir wissen jetzt, dass das meiste, was in diesem Raum lebt, Rauschen ist.

In einer gemeinsamen Regression überleben beide Komponenten: Struktur $\beta = 0,63$ ($t = 17,6$), Aktualität $\beta = 0,27$ ($t = 10,8$). Aktualität ist nicht wertlos. Sie ist etwa ein Viertel dessen wert, was Struktur wert ist, und sollte entsprechend gewichtet werden, statt die Entscheidung zu treiben.

Die praktische Regel: Einen frischen Messwert um rund 90 % nach unten korrigieren, bevor man ihm glaubt. Einen strukturellen Edge um rund 10 %. Die beiden Zahlen unterscheiden sich um eine Größenordnung, und sie gleich zu behandeln ist der teuerste Fehler in diesen Daten.

8. Grenzen

- **Das Zell-Universum ist unsere eigene Setup-Familie** — 23 Setups, über Jahre ausgewählt, weil sie plausibel wirkten. Die Persistenz-Befunde sind interne Vergleiche und davon weitgehend unberührt, die absoluten Niveaus aber optimistisch. Dies ist keine Aussage über alle denkbaren Muster.
- **Ein Datensatz.** Jeder Out-of-Sample-Block hier stammt aus denselben 11,4 Jahren. Wie in jeder Studie, die wir veröffentlichen: Live-Forward ist das einzige echte Out-of-Sample.
- **Halbjahre sind eine Setzung.** Jahres-Buckets ergeben dasselbe Bild (strukturelles Signal 95 %, Auffälligkeit 26 %), die genaue Zerfallsform hängt aber von der Bucket-Länge ab.
- **R ist fat-tailed.** Zell-Perioden mit weniger als 25 Trades wurden ausgeschlossen; Ergebnisse mit n nahe dieser Grenze bleiben rauschig.
- **Nicht getestet:** ob Auffälligkeiten sich in bestimmten Regimes anders verhalten (hohe Volatilität, Event-Tage). Die Zerlegung behandelt alle Perioden gleich.

Die Ein-Satz-Fassung

Ein struktureller Edge überlebt fünf Jahre. Eine frische Auffälligkeit überlebt keine sechs Monate. Auf dem Bildschirm sehen beide identisch aus — die beobachtete Streuung unterscheidet sich um weniger als 10 %. Sie auseinanderzuhalten ist keine Frage von Erfahrung oder Gespür. Es braucht die Rechnung.

Rein statistische Auswertung historischer Marktdaten. Keine Trading-Empfehlung, keine Finanzberatung. Historische Performance ist keine Garantie für die Zukunft. Trading mit Hebelprodukten kann zum Totalverlust führen.