

Der Volumen-Filter, der einen Edge erfand:
Phantom-Edges in volumenarmen Zeitfenstern

Ein Breakout-Setup in der US-Mittagsflaute zeigte +0,90 R bei 60 % Trefferquote — es wäre das beste Setup unseres Systems gewesen. Ohne den Volumen-Bestätigungsfiler: −0,11 R bei 29 %. Der Filter hatte den Trigger im Median um 100 Minuten verschoben, während der Einstieg rechnerisch am alten Level blieb. Ein Lehrstück über eine Lookahead-Falle, die jeder Backtest mit Volumen-Bestätigung reproduzieren kann.

Dokument-Metadaten

Feld	Wert
Herausgeber	KruegerAlgorithms
Titel	Der Volumen-Filter, der einen Edge erfand: Phantom-Edges in volumenarmen Zeitfenstern
Erscheinungsdatum	19. August 2026
Version	v1.0
Verdikt	Referenz
Datengrundlage	Datenbasis: NQ, DOW und SPX, 05.01.2015 – 05.06.2026 , 10-Minuten-Referenzkerze um 18:10 Berlin (US-Mittag), Break in beide Richtungen, Stop an der Gegenseite der Referenzkerze, erster Break zählt, Einstiegs-Cutoff 3 Stunden. Exit: Trailing-Stop BE 0,5 / TS 1,0 / Schritt 0,5, netto nach Spread und Slippage , Standardfehler tages-geclustert. Lag- und Gap-Messung auf NQ und DOW. Keine Handelsempfehlung.
Hinweise	Korrektur vom 29. September 2026; Neurechnung vom 5. Oktober 2026. Wortlaut in der Zusammenfassung.
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

Korrektur vom 29. September 2026: Abschnitt 2 sagt, am Vormittag seien Bruchkerze und Volumenkerze dieselbe, und Abschnitt 4 folgert, die etablierten Vormittags-Setups seien praktisch nicht betroffen. Beides war falsch. Ein Test auf einem Zufallsmarkt zeigte, dass beim stärksten Vormittags-Setup in 39 bis 40 % der Trades das Level schon vor der Volumenkerze gebrochen war; ohne Volumenbedingung fällt sein Ergebnis von +0,60 auf +0,01 R. Der hier beschriebene Mechanismus ist bestätigt, seine Reichweite war größer. Alle R-Werte dieser Studie stammen aus unserer älteren Backtest-Kette. Details: Der Zufallsmarkt als Lineal.

Neurechnung vom 5. Oktober 2026: Ohne den fehlerhaften Einstieg liegen die Mittagsfenster auf der korrigierten Kette bei $-0,10$ bis $-0,17$ R je Trade (alte Kette: $+0,57$ bis $+0,92$ R). Alle Zahlen: Unsere eigenen Ausbruchsregeln, neu gemessen.

Ein Breakout-Setup in der US-Mittagsflaute zeigte $+0,90$ R bei 60 % Trefferquote — es wäre das beste Setup unseres Systems gewesen. Ohne den Volumen-Bestätigungsfiler: $-0,11$ R bei 29 %. Der Filter hatte den Trigger im Median um 100 Minuten verschoben, während der Einstieg rechnerisch am alten Level blieb. Ein Lehrstück über eine Lookahead-Falle, die jeder Backtest mit Volumen-Bestätigung reproduzieren kann.

Volumen-Bestätigung gehört zu den Regeln, die in fast jedem Breakout-Lehrbuch stehen: Ein Ausbruch zählt erst, wenn er von überdurchschnittlichem Volumen begleitet wird. Die Logik ist einleuchtend — ein Break ohne Beteiligung ist ein Fehlausbruch in Wartestellung. Entsprechend selbstverständlich saß eine solche Regel seit Jahren in unserer Signal-Erzeugung: Ein Break gilt erst in der ersten Kerze, die *beides* erfüllt, Level gebrochen **und** Volumen mindestens auf Höhe seines 20-Perioden-Durchschnitts.

Bei unseren etablierten Vormittags-Setups war das nie ein Problem. Dort liegt das Volumen ohnehin über dem Durchschnitt, die Bedingung ist in der Break-Kerze selbst erfüllt, und die Kette funktioniert wie gedacht. Das Problem tauchte erst auf, als wir die Mechanik in ein Zeitfenster trugen, für das sie nie gebaut war.

Diese Auswertung enthält keine Trading-Empfehlung.

Inhaltsverzeichnis

Zusammenfassung	2
Inhaltsverzeichnis	3
Abbildungen	3
1. Der Fund, der zu gut war	4
2. Die Mechanik des Fehlers	5
3. Die Diagnose-Signatur: das Uhrzeit-Raster	5
4. Vier Pflichtchecks für jeden neuen Setup-Test	6
5. Was daraus folgt	6
6. Grenzen	6

Abbildungen

Abb. 1: NQ, 10-Minuten-Kerzen: der Bruch ohne Gate kommt 20 Minuten nach der Referenzkerze, der Bruch mit Volumen-Gate erst 100 Minuten später

Abb. 2: Das Gate hat den Edge erfunden: +0,90 R mit, -0,11 R ohne

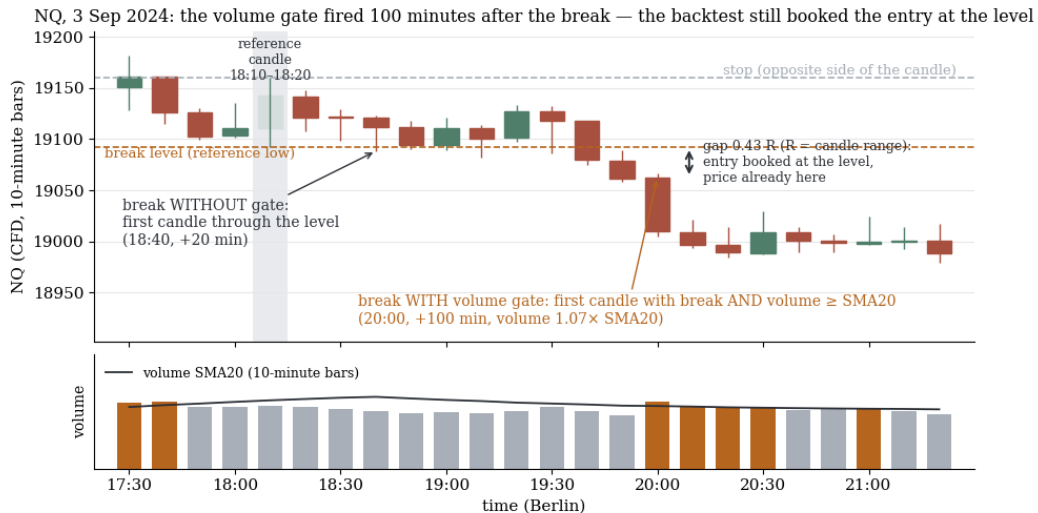


Abb. 1: NQ, 10-Minuten-Kerzen: der Bruch ohne Gate kommt 20 Minuten nach der Referenzkerze, der Bruch mit Volumen-Gate erst 100 Minuten später. Der Backtest bucht den Einstieg trotzdem am Level, obwohl der Kurs längst weiter ist.

1. Der Fund, der zu gut war

The gate invented the edge: $+0.90$ R with the volume gate, -0.11 R without — and the 'edge' tracks the volume lull, not the market

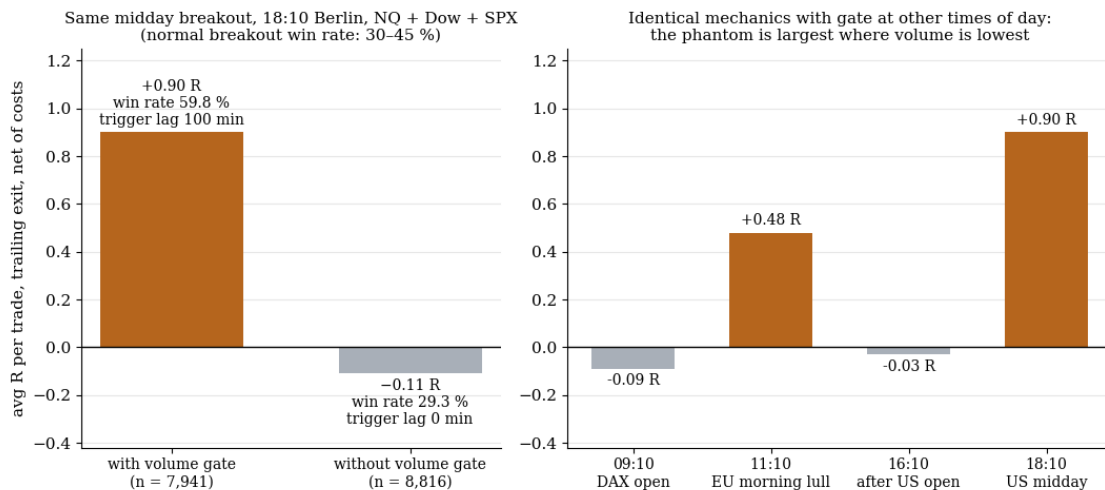


Abb. 2: Das Gate hat den Edge erfunden: $+0,90$ R mit, $-0,11$ R ohne. Dieselbe Mechanik zu anderen Uhrzeiten zeigt, dass der Phantom-Edge dort am größten ist, wo das Volumen am niedrigsten liegt.

Wir testeten ein Referenzkerzen-Breakout in der US-Mittagsflaute — 18:10 Berlin, 12:10 New York, die ruhigste Phase der US-Session. Das Ergebnis:

Variante	n	avgR	Trefferquote	Trigger-Lag (Median)
Mit Volumen-Gate	7.941	$+0,90$ R	59,8 %	100 Min
Ohne Volumen-Gate	8.816	$-0,11$ R	29,3 %	0 Min

Die erste Zeile wäre mit Abstand das beste Setup in unserem System gewesen, jedes einzelne Jahr positiv. Die zweite Zeile ist die Wahrheit. Der komplette „Edge“ von einem R pro Trade war ein Artefakt der Rechenkette.

Die erste Warnung hätte die Trefferquote sein müssen. Ein Breakout mit Stop an der Gegenseite der Referenzkerze liegt bei uns über alle Märkte und Jahre bei **30 bis 45 % Trefferquote**. Das ist keine Schwäche, sondern die Struktur des Setups: Man zahlt viele kleine Verluste für wenige große Gewinne. **60 % Trefferquote**

bei einem Breakout ist kein Fund, sondern ein Bug. Wir haben das nicht sofort gesehen, weil das Ergebnis Jahr für Jahr stabil war — Stabilität ist ein schlechter Bug-Detektor, wenn der Bug selbst stabil ist.

2. Die Mechanik des Fehlers

Zwei Schritte der Kette passten nicht zusammen. Die Signal-Erzeugung setzt den Trigger auf die erste Kerze, die Level *und* Volumenbedingung erfüllt. Der Episoden-Bau setzt den Einstieg unverändert auf Level plus Kosten — so, als hätte man exakt am Break gekauft.

Im Vormittag ist das dieselbe Kerze. Im Mittagsfenster nicht. Dort ist das Volumen fast durchgehend unter seinem 20-Perioden-Durchschnitt, der Preis läuft durch das Level, und die Volumenbedingung wird erst Kerzen oder Stunden später erfüllt — wenn die Nachmittags-Aktivität einsetzt. Zu diesem Zeitpunkt steht der Preis längst weit jenseits des Levels. Der Backtest kauft rückwirkend zu einem Preis, den es nicht mehr gibt.

Gemessen auf NQ und DOW:

Kennzahl	Mittags-Setup (18:10)	Vergleich: Vormittags-Setup
Trigger-Lag (Median)	100 Minuten	0 Minuten
Preis beim Trigger jenseits des Einstiegslevels (Median)	+0,35 R	—
Anteil Trades mit Preis jenseits des Levels	63 %	—
75. Perzentil des Abstands	+1,2 R	—

Der Median-Trade wurde also mit einem geschenkten Vorsprung von 0,35 R gebucht, jeder vierte mit mehr als 1,2 R — und der Stop lag entsprechend weiter weg als real möglich. Daraus entstehen 60 % Trefferquote und +0,90 R, ohne dass irgendein Marktverhalten dahintersteht.

Das ist ein Lookahead-Bias, auch wenn er nicht so aussieht. Kein Feature schaut in die Zukunft; der Einstiegspreis tut es. Die Fehlerklasse ist dieselbe wie in unserem Post-mortem zum Feature-Leck: Ein Wert wird zu einem Zeitpunkt verwendet, zu dem er real nicht mehr verfügbar ist.

3. Die Diagnose-Signatur: das Uhrzeit-Raster

Bevor wir die Ursache im Code hatten, hat uns eine Benchmark gezeigt, dass etwas nicht stimmt. Wir haben die *identische* Mechanik — 10-Minuten-Referenzkerze, Break beide Seiten, Stop an der Gegenseite, Trailing-Exit — auf ein Raster von Uhrzeiten über den ganzen Handelstag gelegt, NQ, DOW und SPX gepoolt:

Uhrzeit (Berlin)	Marktphase	avgR mit Volumen-Gate
09:10	DAX-Open	−0,09 R
11:10	europäische Vormittagsflaute	+0,48 R
16:10	kurz nach US-Open	−0,03 R
18:10	US-Mittag	+0,90 R

Das Muster ist eindeutig: **Der Phantom-Edge ist antikorreliert zur Handelsaktivität.** Genau dort, wo am wenigsten gehandelt wird, ist er am größten; wo der Markt aktiv ist — DAX-Open, US-Open —, verschwindet er. Ein echtes Zeitfenster-Setup würde sich umgekehrt verhalten oder zumindest nicht so sauber der Volumenkurve folgen. Ein Edge, der exakt die Form der Volumenlücke hat, ist ein Edge des Volumenfilters, nicht des Marktes.

Das Raster ist damit ein allgemeines Werkzeug: Wenn ein neues Zeitfenster auffällig gut aussieht, dieselbe Mechanik auf alle anderen Uhrzeiten legen. Verläuft die Güte entlang der Aktivitätskurve, ist die Mechanik der Verdächtige.

4. Vier Pflichtchecks für jeden neuen Setup-Test

Aus diesem Fall haben wir vier Prüfungen abgeleitet, die seitdem jeder neue Setup-Test durchlaufen muss, bevor sein Ergebnis zählt:

1. **Trigger-Lag ausgeben:** Zeit zwischen dem Moment, in dem das Setup bereit ist, und dem tatsächlichen Trigger. Median über 10 Minuten ist Alarm.
2. **Abstand Preis-zu-Level beim Trigger** in R messen, vorzeichenrichtig in Trade-Richtung. Median über null bedeutet: der Einstieg ist fiktiv.
3. **Uhrzeit-Raster als Benchmark**, sobald ein neues Zeitfenster besser aussieht als die etablierten.
4. **Trefferquoten-Plausibilität:** Breakouts mit Stop an der Gegenseite liegen bei 30 bis 45 %. Deutlich mehr ist ein Grund, die Kette zu prüfen, nicht zu feiern.

Der Fix für neue Tests ist trivial — Volumenbedingung aus dem Trigger entfernen, der Break zählt in der ersten Kerze durch das Level, das Lag fällt auf null. Die etablierten Vormittags-Setups waren praktisch nicht betroffen, weil dort das Lag ohnehin bei null lag; ihre Zahlen bleiben gültig. Betroffen war ausschließlich jeder Test, der die Kette in ein volumenarmes Fenster getragen hatte.

5. Was daraus folgt

Ein Filter, der den Trigger verschiebt, verschiebt auch den Einstieg. Jede Bestätigungsregel — Volumen, Schlusskurs-Bestätigung, zweite Kerze — verlegt den Trigger nach hinten. Der Backtest muss den Preis *zum Trigger-Zeitpunkt* verwenden, nicht den Preis am Level. Sobald beides auseinanderfällt, entsteht ein Edge aus nichts.

Zu gute Ergebnisse haben eine Basisrate. Ein Setup, das doppelt so gut ist wie alles, was man bisher gemessen hat, ist mit höherer Wahrscheinlichkeit ein Bug als ein Fund. Wir hatten mit unserer Setup-Familie über elf Jahre eine Vorstellung davon, wie gut ein Breakout sein kann — und haben sie beim ersten spektakulären Ergebnis ignoriert.

Stabilität schützt nicht. Der Phantom-Edge war jedes Jahr positiv, weil die Volumenlücke jedes Jahr an derselben Stelle liegt. Ein Bug, der aus der Marktstruktur folgt, sieht in jeder Zerlegung aus wie ein struktureller Edge.

6. Grenzen

- **In-Sample.** Beide Varianten (mit und ohne Gate) sind auf dem vollen Zeitraum gemessen; da das Ergebnis ein Bug-Nachweis und kein Setup-Fund ist, ist ein OOS-Split hier nicht der Punkt.
- **Die Trefferquoten-Basisrate von 30–45 % gilt für unsere Breakout-Mechanik** mit Stop an der Gegenseite der Referenzkerze. Andere Stop-Konzepte haben andere Basisraten; die Regel ist „kenne deine Basisrate“, nicht „60 % ist immer falsch“.
- **Lag und Gap wurden nur auf NQ und DOW gemessen**, das Uhrzeit-Raster auf NQ, DOW und SPX gepoolt. Auf europäischen Indizes und FX wurde die Größe des Effekts nicht separat quantifiziert.
- **Nicht getestet:** ob eine Volumenbestätigung, *korrekt* mit dem Preis zum Trigger-Zeitpunkt gerechnet, einen eigenen positiven oder negativen Beitrag hat. Diese Studie zeigt nur, dass die falsche Verrechnung einen Phantom-Edge erzeugt — nicht, dass Volumen als Filter wertlos wäre.

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.