The Double Bottom Under the Microscope

A six-market investigation of whether a famous reversal pattern actually predicts price — and why its apparent edge is an illusion.

Research Dataset: Nasdaq-100 (NQ), S&P 500 (ES), Gold (GC), Silver (SI), Crude Oil (CL), and Copper (HG) Futures · hourly bars, April 2019 – May 2026

Author: Dhaval Barot · Published by MPM Markets – Market Probability Model · mpmmarkets.com

Abstract

We test the widely-taught belief that a double-bottom pattern signals a reliable price reversal, using roughly seven years of hourly data across six liquid futures markets and **13,704 detected occurrences**. We measure not only whether the pattern is followed by higher prices, but whether it does so more than a random entry at the same moment — the only test that separates a genuine signal from the rising tide of a bull market. We find that the double bottom carries **no statistically reliable predictive edge**. None of three entry points (formation, confirmation, neckline break) beats random consistently across markets, and none of five traditional quality filters (symmetry, volume, prior trend, low-match tightness, and a purpose-built similarity score) carries forward information. The one effect that appears overwhelmingly significant — the neckline break — proves mechanical: it measures a move already in progress and vanishes when measured forward from a point a trader could act on. The result is consistent across all six instruments. We report the descriptive analysis and the negative tradability result together, with full sample sizes and methodology, so that other researchers can verify, extend, or refute this work.

1. Motivation

The double bottom — a “W” shape where price falls to a low, recovers, falls again to a similar low, then rises — is among the most widely taught reversal patterns in technical analysis. It appears in virtually every trading textbook and course, almost always presented as a reliable signal that a downtrend is ending. Yet that confidence rests on remarkably little evidence: most presentations show a handful of hand-picked successful examples and never address how often it works, whether it beats random entry, or whether the “quality” rules traders apply genuinely help.

We set out to measure the behaviour directly and — separately — to test honestly whether any measured tendency translates into an exploitable edge. We treat these as two different questions, because a statistically real pattern and a profitable strategy are not the same thing. MPM Markets operates on a principle of **reaction, not prediction**: we test whether a setup carries genuine forward information before treating it as tradeable.

Research Snapshot

Research question	**Can the double bottom predict price?**
Markets tested	**6**
Years of data	**7**

Patterns analysed	**13,704**
Timeframe	**Hourly (1H)**
Entry methods tested	**3**
Quality filters tested	**5**
Main result	**No statistically reliable predictive edge**

2. Data and Definitions

Instruments. Six CME-group futures spanning three asset classes: Nasdaq-100 (NQ) and S&P 500 (ES) equity indices; gold (GC), silver (SI), copper (HG) metals; crude oil (CL) energy. Chosen to cross asset classes so that any genuine effect would replicate across unrelated markets, while any market-specific fluke would stand out as isolated.

Sample. 20 April 2019 through 24 May 2026, hourly resolution. Data collected and normalized by the author; timezone handling verified per file.

Pattern definition (deliberately flexible). Rather than a textbook "two equal lows" rule, we detect a family of double-bottom-like structures: two confirmed swing lows (each the lowest point within four bars on either side), whose prices match within the greater of 0.25×ATR or a small fixed points threshold; separated by 8 to 120 hourly bars; with a middle peak (neckline) at least 0.5×ATR above the higher low. This permissiveness lets the data — not our prior beliefs — decide what matters.

Significance. Every effect is tested against a randomized benchmark (entries placed at random in the same market and period, held the same duration). An effect is graded **SIGNAL** if it beats ≥97.5% of random draws, **LEAN** if 84–97.5%, and drift if below. Because we tested roughly thirty-five market-by-question combinations, we trust a finding only if it reaches SIGNAL *and* replicates across multiple markets — a lone significant cell is treated as noise.

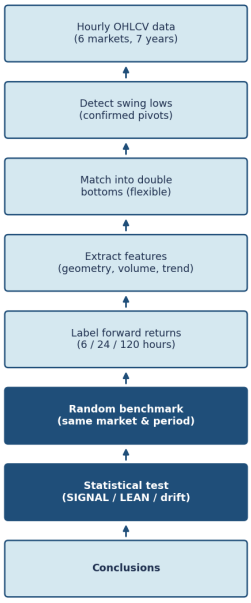

Figure 1. The research pipeline: each pattern flows from raw data through detection, measurement and labelling to a randomized statistical test before any conclusion is drawn.

3. Principal Finding: The Pattern Does Not Predict

Across all 13,704 patterns, raw forward outcomes were 52.2% positive at 6 hours, 54.0% at 24 hours, and 57.8% at 120 hours. Taken naively this looks bullish — but it is exactly what market drift produces in a rising market, and establishes nothing on its own. The randomized benchmark is what separates signal from drift.

3.1 No quality filter carries information

Each filter compares high-quality against low-quality patterns *within the same period*, a comparison drift cannot distort. Pooled across all markets (n = 13,704), 24-hour horizon:

Question (pooled)	High	Low	Gap	Beats	z	Verdict
Q1 Similarity (top 25% vs rest)	2.75	1.71	1.04	76.8%	0.73	drift
Q2 Prior trend down vs other	1.07	2.73	-1.66	8.8%	-1.33	drift
Q3 Volume high vs low	2.29	1.63	0.66	69.3%	0.51	drift
Q4 Neckline broke vs not	20.26	-11.20	31.46	100%	24.65	**SIGNAL**
Q5 Tight vs loose lows	2.53	1.40	1.13	81.5%	0.90	drift

Every quality filter is drift except the neckline break (examined separately below, because it is mechanical). Per-market results confirm the pooled picture, with only scattered isolated LEANs of the kind chance guarantees. Notably, prior trend ran *opposite* to textbook — patterns after downtrends did marginally worse — and volume was significantly *negative* in gold. None of the traditional "quality" beliefs survived.

3.2 The similarity score has no predictive gradient

We designed a 0–100 similarity score to capture everything that should make a double bottom "cleaner." Sorted into deciles (pooled, 24-hour return in points), the relationship is non-monotonic noise:

Decile (low → high)	1	2	3	4	5	6	7	8	9	10
Mean return (pts)	1.17	2.39	1.39	1.76	-0.43	-0.76	5.89	4.04	0.48	3.70

The highest-scoring decile (3.70) is beaten by the seventh (5.89) and barely differs from the second; the middle deciles are negative. The intuition that a prettier pattern is a better pattern is, on this evidence, simply false.

4. The Neckline-Break Illusion

One effect appeared overwhelming. Patterns breaking the neckline within 24 hours outperformed those that did not — **SIGNAL** in all six markets simultaneously, with extreme margins (z = 17–25):

Market	Broke (pts)	Did not	Gap	z	Verdict
ES	23.56	-16.22	39.78	18.37	**SIGNAL**
NQ	107.30	-67.81	175.11	17.71	**SIGNAL**
GC	15.26	-6.85	22.12	18.19	**SIGNAL**

SI	0.34	-0.16	0.51	18.55	**SIGNAL**
CL	0.86	-0.53	1.39	20.71	**SIGNAL**
HG	—	—	—	20.60	**SIGNAL**

The central distinction of this paper

This is an artifact, not a signal. The measurement window *included the break move itself*. Saying "patterns that broke upward then showed higher returns" is close to saying "things that went up, went up." It is mechanically true and untradeable, because by the time the break is visible, the move it measures has already occurred. A real pattern can describe price action without predicting it — and the test below confirms exactly this.

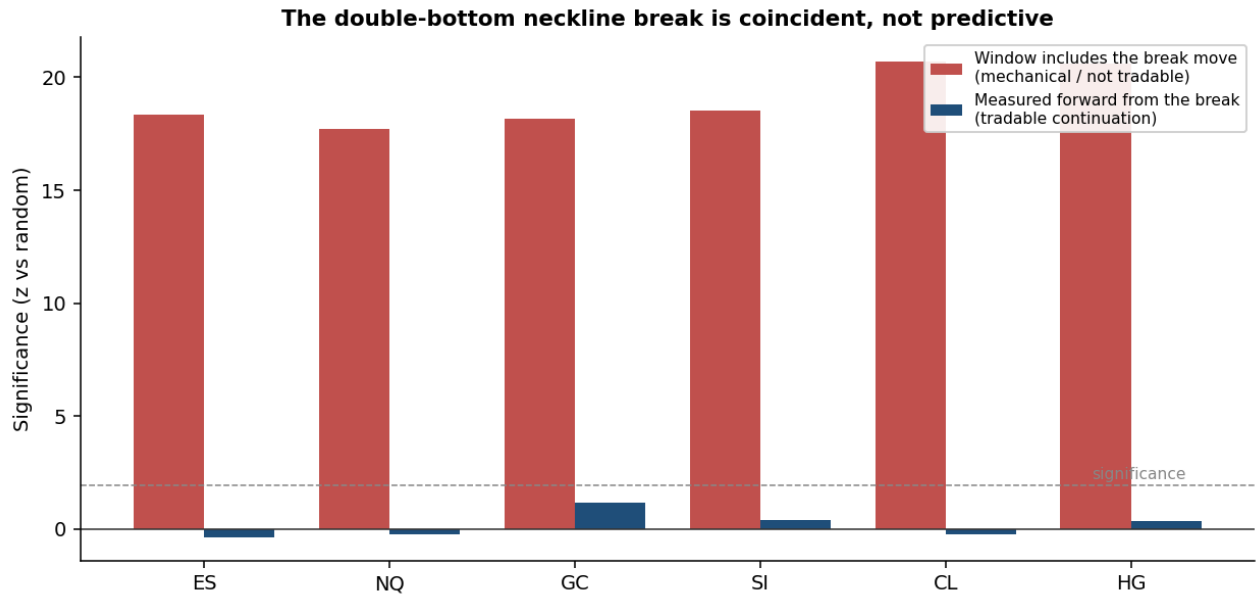

Figure 2. The same break effect measured two ways. Red: a window that includes the break move — hugely significant everywhere, but mechanical. Navy: measured forward from the break, the only part a trader could capture — it collapses to near zero.

5. Tradability Testing: Three Entry Points, All Fail

To be sure the pattern was untradeable rather than merely mis-measured, we tested three distinct entry points, each benchmarked against random entries in the same period. The formation and break tests **included patterns that went on to fail** (broke downward), so no survivorship bias could flatter the results.

5.1 Post-break continuation (enter on the break)

Entering at the break bar and measuring forward — the only part a trader can capture:

Market	H = 6h	H = 24h	H = 120h
ES	drift (z 0.18)	drift (z -0.34)	drift (z -0.48)
NQ	**LEAN (z 1.34)**	drift (z -0.21)	drift (z -0.78)
GC	**LEAN (z 1.60)**	**LEAN (z 1.18)**	drift (z -0.04)

SI	drift (z 0.77)	drift (z 0.40)	drift (z 0.47)
CL	drift (z 0.67)	drift (z -0.23)	drift (z 0.03)
HG	drift (z 0.33)	drift (z 0.36)	**SIGNAL (z 2.38)**

No horizon shows multi-market SIGNAL (counts: 0/2/4, 0/1/5, 1/0/5). The lone SIGNAL (HG, 120h) is isolated with its shorter horizons flat — the textbook chance result. The break does not lead the move.

5.2 Formation entry (buy the W as it forms, including failures)

The earliest entry: buy as the second low confirms, before any break, including patterns that later fail:

Market	H = 6h	H = 24h	H = 120h
ES	drift (z 0.83)	drift (z 0.10)	drift (z 0.55)
NQ	drift (z 0.36)	drift (z -0.43)	drift (z 0.18)
GC	drift (z 0.33)	drift (z -0.09)	drift (z -0.77)
SI	drift (z -0.10)	drift (z 0.06)	drift (z 0.53)
CL	**LEAN (z 1.40)**	**LEAN (z 1.23)**	**LEAN (z 1.26)**
HG	drift (z -0.15)	drift (z 1.04)	**SIGNAL (z 2.91)**

Again no multi-market signal (counts: 0/1/5, 0/1/5, 1/1/4). Crude oil shows weak LEANs across all three horizons — the most coherent flicker in the study — but it is one market of six, never reaches SIGNAL, and involves tiny per-trade magnitudes. The isolated HG-120h cell reappears, reinforcing that it is a quirk of copper's slow drift, not a double-bottom effect.

6. Conclusions

> **Key findings at a glance**
>
> ✓ Tested 13,704 patterns across six futures markets · three entry methods · five quality filters
> ✓ Every effect benchmarked against randomized entries (drift-controlled)
> ✗ **No statistically reliable predictive edge at any entry point**
> ✗ No quality filter — including a purpose-built similarity score — carried information
> ○ The neckline "break" effect proved mechanical (coincident), not predictive

The double bottom is descriptive, not predictive. Across six futures markets and 13,704 occurrences, it does not forecast price at any tested entry point, and none of its traditional quality filters add information. The neckline break, the one effect that appears significant, is coincident with the move rather than leading it — a confirmation of price action already underway, not a prediction of price action to come.

7. Practical Takeaway

For a practitioner, the useful conclusion is informational, not directional. The double bottom is a legitimate way to *read* what the market has done — the neckline is a real reference level around which price reacts — but it is not a reason, by itself, to enter a trade. Don't trade a setup because it is a famous pattern; verify every belief

statistically; and beware confirmation that arrives too late. Anyone trading double bottoms profitably is most likely being helped by some other factor they have not isolated. The value of this finding is in understanding market behaviour, not in a ready-made trade.

8. Limitations

- **Bull-market regime only.** 2019–2026 rose overall. The pattern's behaviour in a sustained bear market is untested; a negative result here is strong but not a universal law.
- **Hourly timeframe and one detection rule.** A different timeframe or a materially different definition of "double bottom" could yield different results, though the flexible definition was chosen to minimise this.
- **Finite entry space.** We tested three entry points; other exit models, holding periods, or regime filters could in principle differ. Our negative result applies to what was tested.
- **Hypothetical.** All results are derived from historical data with modelled execution; they are not live trading records.

9. What a Researcher Could Try Next

- Different market regimes — particularly a sustained bear market, the one condition this study could not test.
- Different timeframes — daily and intraday-minute structures may behave differently from hourly.
- Adaptive, volatility-scaled tolerances for pattern detection.
- Intraday session effects — whether patterns completing in specific sessions (e.g. the US open) differ.
- The crude-oil flicker — the one weakly-consistent cross-horizon result, worth a dedicated, pre-registered test before any weight is placed on it.

Why we publish a study with no edge

Most trading research shared publicly reports only successes. We believe rigorous research should also report findings that fail to produce an edge. Publishing negative and null results reduces survivorship bias, guards against overconfidence, and gives a more honest picture of how markets actually behave. We would rather show our work — including where it did not lead to a tradable strategy — than present a flattering result we could not stand behind. This is the essence of **reaction, not prediction**: we do not assume a setup works because it is famous; we test it, report what we find, and trust only what survives.

10. Reproducibility

Parameter	Value
Pivot lookback (swing confirmation)	4 bars each side
ATR length	14
Low-match tolerance	max(0.25 × ATR, small fixed points per market)
Min / max separation between lows	8 / 120 hourly bars
Min neckline depth	0.5 × ATR above the higher low

Forward horizons	6, 24, 120 hours
Random simulations per test	1,000–5,000
Total patterns	13,704 (ES 1,841; NQ 1,671; GC 2,355; SI 2,723; HG 2,404; CL 2,710)

About the research. This study was conducted and authored by Dhaval Barot and published by MPM Markets – Market Probability Model. Questions, replication requests, and research discussions may be submitted through the contact form at mpmmarkets.com/contact.

Suggested citation. Barot, D. (2026). *The Double Bottom Under the Microscope: A Six-Market Investigation of Whether a Famous Reversal Pattern Actually Predicts Price.* MPM Markets.

First published: June 2026. Research conducted and authored by Dhaval Barot, MPM Markets.