

Predictability at the Execution Boundary: Threshold-Crossing Dynamics and the Capacity Ceiling in Taiwan Index Futures*

Yi-Hao Lai[†]

Version 1 — 22 August 2026

Abstract

We measure the intraday predictability of threshold-crossing sequences in Taiwan index futures using the complete tick record of 3,575 trading sessions (207.1 million day-session trades, 2012–2026), and show that the predictability is real, large under the right conditioning, and yet unavailable to a liquidity-taking trader for two distinct and separately measurable reasons.

First, execution absorbs 90–107% of the gross edge across thresholds spanning two orders of magnitude (1 to 160 index points). Reversals travel farther than continuations—the asymmetry grows from 0.479 to 5.257 points as the threshold widens from 5 to 40—so the true gross edge is $1.7\times$ to $5.3\times$ larger than a symmetric-payoff approximation implies. But with next-tick fills the realised gross P&L is

***Working paper. Not under submission, and no submission is planned.** The manuscript is kept as a citable, reproducible record of the measurement; the reasons it is shelved are recorded in the accompanying `README.md`. Evidence base: experiment K1824, 207,105,249 day-session trades of TAIEX TX index futures, 2012-01-02 to 2026-08-20; every table is reproducible from the committed result JSONs. The preregistration declares `is_blind = false`: the motivating counts were observed before the analysis plan was frozen, and any citation of these results must carry that qualification.

[†]Department of Finance, Da-Yeh University. Email: yhlai@mail.dyu.edu.tw

−0.016 to +0.326 points, leaving nothing for the proportional transaction tax to consume.

Second, the conditional edge is capacity-bound. Conditioning on incident momentum—the size of the move that produced the crossing, relative to the threshold—raises the reversal probability from an unconditional 50.8–54.4% to 96%. That is the strongest directional signal in the study. However, the median fill available on the entry tick is **2 lots**, and 88.8–95.8% of fills are 5 lots or fewer. The edge appears precisely when liquidity has disappeared, which is structural rather than incidental: the trade supplies liquidity, and its return is the compensation for supplying it.

We further decompose the reversal rate into a bid-ask bounce component and a genuine reversion component, and show they behave differently over fourteen years. The one-step hazard $h(1)$ has no trend: the era averages are 55.91%, 56.35% and 56.61% for 2012–2016, 2017–2021 and 2022–2026, and the annual series oscillates between 52.58% and 57.92%. The multi-step hazard $h(k \geq 2)$ declines monotonically over the same period, from 55.67% to 50.63%. The contrast is what survives; attributing the absence of trend in $h(1)$ to the fixed one-point tick size was tested directly and withdrawn (Section 6, claim 4).

Contribution. The paper locates the Grossman–Stiglitz (1980) equilibrium empirically, in a market where the transaction tax is proportional to contract value and therefore quadrupled in point terms as the index level rose $5.5\times$. The ratio of gross edge to execution loss is pinned between 0.66 and 1.07 across a 160-fold range of thresholds: the predictability is worth almost exactly what it costs to extract.

Keywords: market microstructure, intraday predictability, execution cost, capacity constraint, transaction tax, TAIFEX

JEL Classification: G12, G14, G17

1 Introduction

A price series crossing a fixed threshold generates a sequence of events whose statistical structure has been studied since Guillaume et al. (1997) and catalogued extensively by Glattfelder, Dupuis and Olsen (2011), who report twelve empirical scaling laws in foreign exchange data. That literature establishes what the events look like. It does not establish what a trader reacting to them would earn, because it is built on FX quote data where the tick is effectively continuous and there is no transaction tax.

This paper asks the trading question directly, on a market chosen because it makes the costs impossible to ignore: TAIFEX TX index futures, where the tick size has been exactly one index point for the entire sample, and where the futures transaction tax is levied at 0.00002 of contract value per side—so the same point-denominated strategy cost roughly 0.53 points per round trip in 2012 and 2.09 points in 2026, purely because the index rose.

The central finding is not that predictability is absent. It is present, it is large under the right conditioning, and we measure it precisely. The finding is that it sits exactly at the execution boundary, and that a second constraint—available size at the moment the signal fires—binds even where the first does not.

The phenomenon itself is not new. Partial reversal after a large move is the transient component of price impact (Bouchaud, Farmer and Lillo 2009), and Vlasiuk and Smirnov (2025) document the same push-response asymmetry in SPY. What is missing from that work, by its own account, is the cost and capacity overlay: they leave an explicit transaction-cost treatment to future work. This paper supplies it, in a market whose one-point tick and proportional transaction tax make both constraints unusually legible (Section 5.1).

2 Data

Five data-handling decisions are load-bearing.

Column access by header name. The file format changes within the sample:

Table 1: Sample. The tick record is complete: no session in the period is missing, and no observation is winsorised.

Item	Value
Instrument	TAIFEX TX index futures, near-month (TX1)
Period	2012-01-02 to 2026-08-20
Sessions	3,575
Day-session trades	207,105,249
Session analysed	08:45:00–13:45:00
Tick size	1 index point (unchanged throughout)
Contract multiplier	NT\$200 per point

2012-01-02 has nine columns and no opening-auction flag; by 2012-12-03 there are ten. All access is by header name; no column index is hard-coded.

Opening call auction excluded. Rows flagged * are batch clearing prices, not continuous trades. 3,466 such prints are removed. This matters because the auction print is exactly where the largest apparent slippage occurs, and including it would inflate every overshoot statistic.

Per-session base reset. Each session begins a fresh crossing chain, so the overnight gap never enters the first intraday move.

Contract rollover cut. The crossing chain is severed when the expiry-month field changes within a session, so a roll does not register as a price move.

Nothing is winsorised. In particular, the 2017-08-03 opening flash crash—615 prints more than 200 points from the day’s median inside the single second 08:45:25, minimum print 9,408 against a daily median of 10,397, largest single jump 843 points, every print 2–4 lots, all in the same contract month—is genuine trade data and is retained. Its influence is measured and reported throughout (Section 6).

3 Method

3.1 Crossing definition, and why it differs from the standard one

We record a crossing whenever the price moves δ or more from the price at which the previous crossing occurred, and reset the reference to the new traded price. Direction is not constrained.

This differs materially from the directional-change (DC) algorithm of Glattfelder et al. (2011), in which δ is measured *from the running extremum* and the mode *alternates on every event*. Under that definition a DC is a reversal by construction: the reversal rate is identically one, and successive same-direction events cannot occur. Their framework therefore admits no reversal rate, no run-length hazard, and no reversal-versus-continuation amplitude comparison. Ours admits all three, at the cost of not being directly comparable to their scaling laws.

We implement both on the same data (Section 4.1) so the relationship is measured rather than asserted.

3.2 Trade construction and execution assumption

Let $d_i \in \{-1, +1\}$ be the direction of crossing i and let P_i be the first traded price *after* crossing i . The fade trade for crossing i enters at P_i in direction $-d_i$ and exits at P_{i+1} :

$$\text{PnL}_i = (P_{i+1} - P_i) (-d_i). \quad (1)$$

Using the next traded price, rather than the crossing price, is the whole point. A taker reacting to a crossing cannot transact at the crossing price; the overshoot is therefore already inside the fill, and no separate slippage term is applied. Adding one would double-count.

3.3 Cost model

Costs are charged per leg at the traded price. With $\tau = 0.00002$ the statutory transaction tax rate on contract value and $f = \text{NT\$}25$ the per-side brokerage estimate, the round-trip cost in index points is

$$c_i = (P_i + P_{i+1}) \tau + \frac{2f}{\text{NT\$}200}. \quad (2)$$

The first term is the statutory tax on contract value (both sides); the second is a brokerage estimate at the low end of the observed market range, which is not an official rate and is treated as a sensitivity parameter. Expressed in points, the tax component is 2τ times

the index level per round trip and is therefore independent of the contract multiplier—identical in points for the large and mini contracts.

4 Results

4.1 Reconciliation with the scaling-law literature

Implementing the Glattfelder et al. DC algorithm on TX gives a count exponent of

$$N(\delta) \propto \delta^E, \quad E = -1.888, \quad (3)$$

against the random-walk value of -2 and their FX estimates near -2 . The implied count ratio per doubling of the threshold is 3.70, not 4.

This is worth stating plainly because it is the origin of the puzzle that motivated the study: an observation that crossing counts fall by roughly one third rather than one quarter when the threshold doubles is **a restatement of a scaling law first reported by Guillaume et al. (1997)**, not a new empirical fact. Under our own crossing definition the ratios are 3.735, 3.767 and 3.855 for the $5 \rightarrow 10$, $10 \rightarrow 20$ and $20 \rightarrow 40$ doublings, implying Hurst exponents of 0.526, 0.523 and 0.514—close to, and converging on, the random-walk value.

Their overshoot law survives on TX. Mean overshoot divided by threshold is 0.896, 0.971, 0.987, 0.953 and 0.900 at $\delta = 5, 10, 20, 40, 80$: approximately one threshold, as in FX.

The definitions nonetheless measure different objects, and the clearest evidence is the share of zero overshoots. Under the DC definition it is 27.1% at $\delta = 5$, falling to 5.8% at $\delta = 80$. Under ours it is **73.9% to 76.6%**, because a single crossing frequently stops exactly on the threshold when the tick is a full point. The FX-derived law is intact; the object it describes is not the object a taker faces.

4.2 The hazard decomposition

Let $h(k)$ be the probability that crossing $k+1$ reverses given k consecutive same-direction crossings. Pooled over the sample, $h(k)$ declines monotonically from 71.9% to 51.2%, which reads as trend self-reinforcement.

That decline is an artefact. It is survivorship under unobserved heterogeneity: pooling a 2012 sub-sample with a 78% reversal rate and a 2026 sub-sample with 51% means long runs are populated almost entirely by the low-reversal regime. Mixing two memoryless sub-samples reproduces the entire pooled curve for $k \geq 2$ to within one percentage point.

Stratified by era, $h(k)$ is flat for $k \geq 2$ (Table 2).

Table 2: Reversal hazard $h(k)$ by era at $\delta = 5$. The eras are indistinguishable at $k = 1$ and separate only for $k \geq 2$.

Era	$h(1)$	$h(2)$	$h(3)$	$h(4)$	$h(5)$	$h(8)$
2012–2016	55.91%	54.80%	54.42%	55.20%	54.63%	53.69%
2017–2021	56.35%	53.65%	53.26%	53.53%	54.29%	52.54%
2022–2026	56.61%	51.81%	50.48%	50.61%	50.47%	49.89%

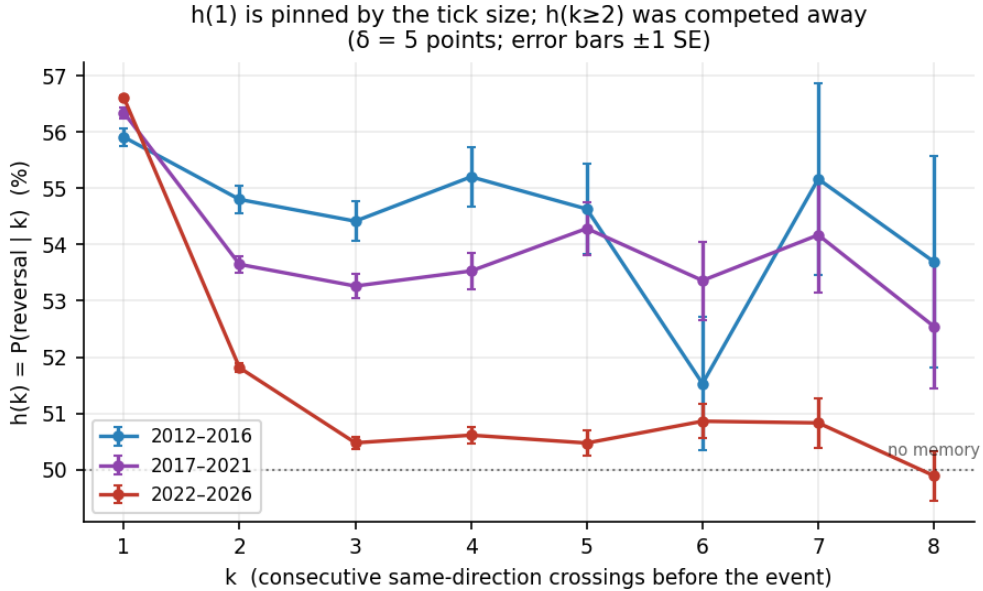


Figure 1: $h(k)$ by era at $\delta = 5$, with ± 1 standard-error bars. The three eras are indistinguishable at $k = 1$ and separate only for $k \geq 2$.

Three consequences follow.

1. **Run length carries no tradable information within a regime.** “It has gone k times in a row, so it is due to turn” has no support.
2. **$h(1)$ has no trend across eras**—55.91, 56.35, 56.61—but the annual series runs from 52.58% (2022) to 57.92% (2026), so the era averages conceal real year-to-year movement. The component is bid-ask bounce; what these numbers do not establish is that it is pinned there by the tick size (Section 6, claim 4).
3. **$h(k \geq 2)$ declines monotonically, from 55.67% to 50.63%.** The genuinely predictable component was competed to nearly zero while the mechanical one oscillated around a flat level.

This decomposition is unavailable under the alternating-mode DC definition, since same-direction runs cannot occur there.

4.3 Amplitude asymmetry and the execution boundary

A symmetric-payoff hurdle, $(2p - 1)\delta > c$, assumes reversals and continuations travel equal distances. They do not (Table 3).

Table 3: Amplitude asymmetry. Reversals travel farther than continuations, and the gap widens with the threshold, so the symmetric-payoff approximation understates the gross edge by a factor of 1.7 to 5.3. All quantities in index points unless marked.

δ	Reversal rate	Move if reversed	Move if continued	Asymmetry	Symmetric edge	True gross edge
5	54.38%	5.985	5.506	+0.479	0.438	0.743
10	51.57%	11.524	10.573	+0.951	0.314	0.822
20	50.91%	22.769	20.616	+2.153	0.365	1.472
40	50.78%	46.132	40.875	+5.257	0.626	3.309

The true gross edge is $1.7\times$ to $5.3\times$ the symmetric approximation. Execution then removes essentially all of it (Table 4).

Execution loss is bounded in absolute size (roughly one to three points) while the edge scales with the threshold, so its share falls monotonically. The apparent profitability at $\delta \geq 80$ is addressed in Section 6.

Table 4: The execution boundary. Entry and exit are the first traded prices after successive crossings, with costs charged per leg at the traded price (equation 2). Execution loss is the true gross edge less the realised gross P&L, expressed as a share of the former.

δ	True gross edge	Realised gross	Execution loss	Cost	Net	t
1	0.599	0.061	89.8%	0.868	-0.807	-4064
5	0.743	-0.016	102.2%	1.344	-1.360	-310
10	0.822	-0.055	106.7%	1.414	-1.469	-92
20	1.472	0.051	96.5%	1.459	-1.408	-23
40	3.309	0.326	90.2%	1.509	-1.184	-5.0
160	40.648	13.769	66.1%	1.648	+12.121	2.76

The positive net P&L at $\delta = 160$ does not survive Section 6: it is negative out-of-sample and negative once the single session 2017-08-03 is removed.

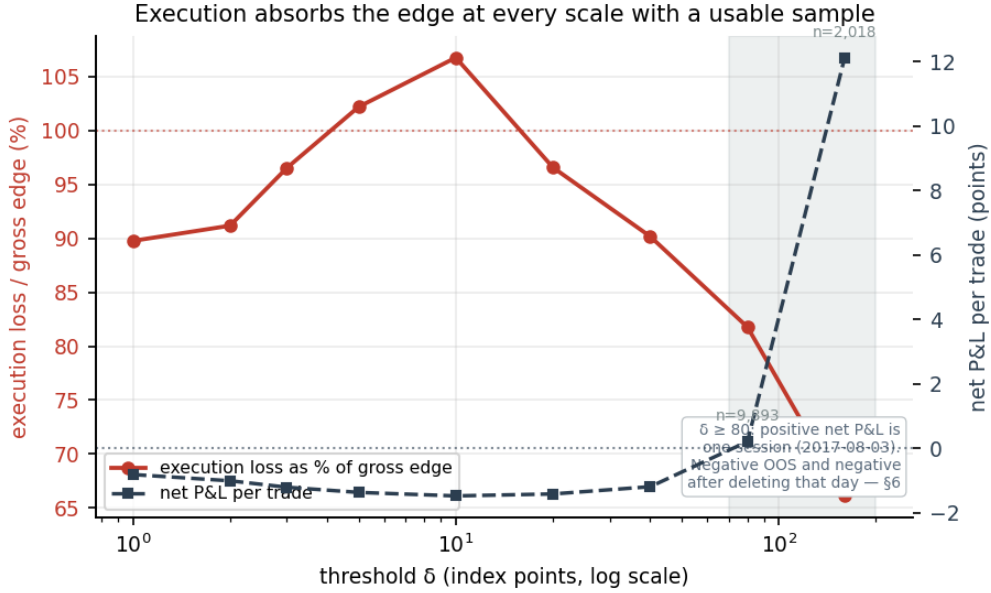


Figure 2: Execution loss as a share of the gross edge (left axis) and net P&L per trade (right axis) against threshold. The shaded region marks $\delta \geq 80$, where net P&L turns positive on a sample driven by one session; Section 6 shows all of those levels are negative out-of-sample and negative once 2017-08-03 is removed.

Volatility-scaled thresholds partially recover the loss. Setting $\delta = k\sigma$, where σ is the previous session’s realised volatility, dominates fixed point counts at every comparable size: -0.627 versus -1.469 at 8–10 points, and -0.561 versus -1.408 at 17–20 points. Execution loss share falls from 106.7% to 82.0%. Roughly 0.8 points is recoverable by matching the threshold to how fast price is moving. Price-proportional (basis-point) thresholds, by contrast, are *worse* than fixed point counts, because the tick size is fixed in points: at 2.5 basis points the 2012 threshold is 1.86 points, under two ticks, and measures almost pure bounce. Part of the apparent cost saving under volatility scaling is a sample-composition effect—more trades land in low-index, low-tax years—and is not counted as improvement.

4.4 Momentum conditioning: a 96% reversal rate

Let r be the size of the move producing the crossing, divided by δ . The reversal rate rises sharply in r (Table 5).

Table 5: Reversal rate conditional on incident momentum r . Against an unconditional 50.8–54.4%, this is the strongest directional structure in the study.

δ	$r \in [1.5, 2)$	$r \in [2, 3)$
5	57.85%	60.03%
20	78.48%	86.88%
40	90.32%	93.27%
80	93.33%	96.36%

Against an unconditional 50.8–54.4%, this is the strongest directional structure in the study: the harder price is struck, the more certainly it rebounds.

4.5 The capacity ceiling

The conditional edge is nonetheless unharvestable, for a reason that is not a cost at all (Table 6).

At $\delta = 40$ with $r \geq 2$ the out-of-sample edge is 20.776 points against a median of 2 available lots: an upper bound of roughly NT\$8,310 per opportunity. **The economic**



Figure 3: Reflection rate against the incident-momentum ratio r , one line per threshold. The signal is directional only: Section 4.5 and Figure 4 show the fill available when it fires has a median of 2 lots.

Table 6: Size available on the entry fill. The median is two lots at every threshold, and it is an upper bound: it assumes the entire print is obtainable and ignores queue priority.

δ	Median lots	Mean lots	Share of fills ≤ 5 lots
5	2	2.46	95.8%
10	2	2.56	94.6%
20	2	2.79	91.7%
40	2	3.05	88.8%

value of this signal is $\text{edge} \times \text{fillable lots} \times \text{frequency}$, not $\text{edge} \times \text{frequency}$, and the first two terms move in opposite directions.

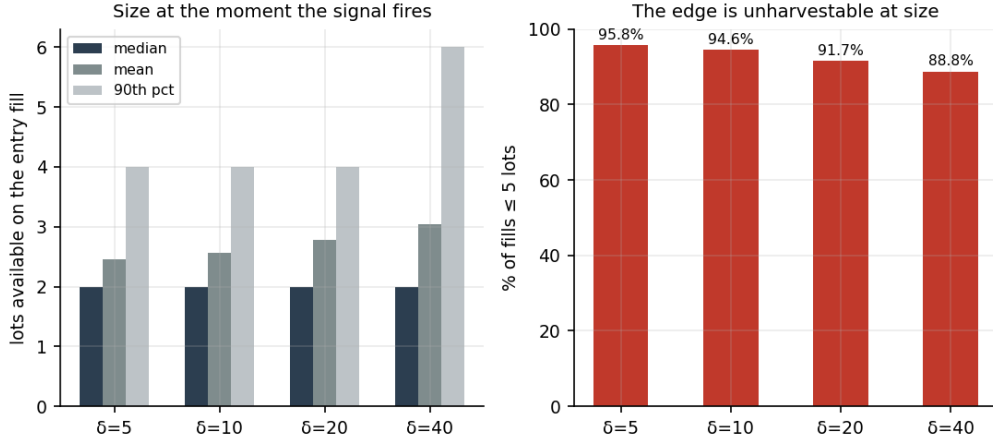


Figure 4: Size available on the entry fill. Left: median, mean and 90th percentile lots by threshold. Right: share of fills at 5 lots or fewer.

This is structural. The trade fades a sweep, supplying liquidity at the moment liquidity has been consumed; the return is the compensation for supplying it, and it is available in the size that was missing. It belongs to the passive side of the book. The trendless $h(1)$ of Section 4.2 is the same fact seen from another angle: bid-ask bounce is market-maker revenue, not taker opportunity.

5 Discussion

The results locate a Grossman–Stiglitz (1980) equilibrium empirically. Predictability is not competed to zero; it is competed to the point where extracting it costs what it is worth. The ratio of realised gross edge to execution loss lies between 0.66 and 1.07 across thresholds spanning 1 to 160 points—a 160-fold range—which is a tighter statement than the usual observation that strategies do not survive costs.

Two features of this market sharpen the test. The tick size has been one point for fourteen years, so the mechanical component of reversal faces an unchanged microstructure while the index level, and hence the proportional tax, rose $5.5\times$. And the transaction tax is levied on contract value, so the cost of a point-denominated strategy quadrupled over the sample without any change in the strategy. The decomposition in Section 4.2

exploits the first; the cost model in Section 3.3 is forced to respect the second.

5.1 What is already known, and what is not

The measurement device here originates in the scaling-law literature, but the economics belong to price impact, and it is worth being explicit about which parts of this paper are new.

Already established. The threshold-crossing / directional-change framework is Guillaume et al. (1997) and Glattfelder et al. (2011); the count scaling law that motivated this study is theirs, not ours (Section 4.1). The transient component of price impact and its decay through passive limit-order counter-flow—“liquidity refill”—is the propagator literature (Bouchaud, Farmer and Lillo 2009). That large moves are followed by partial reversal is therefore an expected consequence of an established model, not a discovery. Most directly, Vlasiuk and Smirnov (2025) document exactly this push-response asymmetry in SPY at high frequency, reporting stronger positive responses after large negative pushes, consistent with asymmetric liquidity replenishment after sell-side impact.

What this paper adds. Vlasiuk and Smirnov state that incorporating an explicit transaction-cost overlay is left to future work. That overlay is this paper’s subject. Specifically:

1. **The execution boundary is quantified, not asserted.** The ratio of gross edge to execution loss is 0.66–1.07 across thresholds from 1 to 160 points. The literature establishes that transient impact decays; it does not establish how much of that decay is reachable by a trader who must transact at the next available price.
2. **Capacity is measured.** The median fill when the signal fires is 2 lots. This converts the economic question from “is there an edge” into “what size does the edge admit”, and the two answers move in opposite directions as the threshold grows.
3. **The bounce component is separately identified.** The two components move differently over fourteen years: $h(1)$ has no trend (era averages $55.91 \rightarrow 56.35 \rightarrow 56.61$, annual range 52.58–57.92%) while $h(k \geq 2)$ declined monotonically from 55.67% to

50.63%. The contrast is what the decomposition delivers. Attributing the absence of trend in $h(1)$ to the fixed one-point tick size is a further step, and our day-versus-night test failed to make it: see Section 6, claim 4.

4. **The cost is proportional, not fixed.** The transaction tax is levied on contract value, so an unchanged point-denominated strategy became four times more expensive over the sample. Fee-free or fixed-fee venues cannot separate cost drift from strategy drift.

The remaining exposure is that items 1 and 2 may exist in the optimal-execution literature in a form we have not located. That literature is large and is oriented toward executing a known parent order rather than toward harvesting a signal, but the boundary is not sharp.

6 Robustness, and four claims withdrawn

The apparent profitability at large thresholds is a single trading session (Table 7).

Table 7: Large-threshold forensics. Net P&L per trade in index points. The best single session is 2017-08-03 at all four thresholds.

δ	Full sample	In-sample	Out-of-sample	Excluding 2017-08-03	Median
40	-1.184	+9.015	-1.810	- 1.894	-3.661
80	+0.216	+61.641	-2.697	- 2.825	-2.096
160	+ 12.121	+154.801	-2.194	- 2.322	-0.626
320	+63.311	+214.785	-4.011	- 2.961	+2.874

The full-sample figure at $\delta = 160$ carries $t = 2.76$; it is the only level whose unconditional mean is significant, and it does not survive either column to its right.

Deleting one session turns all four negative. Out-of-sample is negative for all four. A mechanical gate—out-of-sample positive **and** surviving deletion of the best single day **and** a majority of years positive—now governs this class of claim. It was generalised out of this study into `src/volpred/research/tail_robustness_gate.py` on 2026-08-22 and now applies to every strategy candidate the project evaluates.

The frozen conditioning grid (session $\times$ run length $\times$ crossing duration $\times$ prior-day volatility state) yields 222 evaluable cells: 32 in-sample positive, 8 out-of-sample positive, 2 positive in both. Under independence the expected both-positive count is $222 \times 0.144 \times 0.036 \approx 1.15$. Two survivors are consistent with chance, and both are the same flash-crash cell.

Cross-threshold agreement is **not** independent evidence: the daily P&L correlation of the high- r band across thresholds averages 0.895 (0.997 between $\delta = 10$ and $\delta = 20$) with best-five-day overlap of 4 of 5. One 80-point sweep is $r = 16$ at $\delta = 5$ and $r = 2$ at $\delta = 40$, so agreement across thresholds is one observation counted several times.

Four claims made during the investigation were withdrawn on evidence, and are recorded because the reasons are informative.

1. **“Overshoot is threshold-invariant.”** Withdrawn. Full-sample overshoot is 0.766, 1.059, 1.681 and 3.319 at $\delta = 5, 10, 20, 40$ —rising and accelerating.
2. **“Thresholds should be set in relative (basis-point) units.”** Withdrawn. Basis-point thresholds widen the cross-era gap from 1.36 to 26.75 percentage points, because the tick is fixed in points.
3. **“The coefficient of restitution shows viscoelastic dissipation.”** Withdrawn as a ratio bias. The coefficient $e = m_{\text{out}}/m_{\text{in}}$ has the conditioning variable in its denominator, so the correct null is $m_{\text{out}} \perp m_{\text{in}}$, which mechanically implies $e \propto 1/m_{\text{in}}$. Under that null the top- r group at $\delta = 40$ should show $e \approx 0.54$; the observed 0.93 points the opposite way. The median e of exactly 1.000 is tick discreteness, not elasticity. Reflection rate carries no such bias and is retained.
4. **“The mechanical layer is held fixed by the tick size.”** Withdrawn on our own identification test. The wording rested on three-era averages of $h(1)$ and on a day-versus-night comparison meant to separate a tick effect from a liquidity effect; neither supports it. The annual $h(1)$ series ranges from 52.58% (2022) to 57.92% (2026), so “unchanged” overstates what a three-era average shows. And the day–night gap at $\delta = 5$ is 3.71 percentage points but falls to 0.43 at $\delta = 20$, while the day/night tick-

density ratio is $1.87\times$ at all three thresholds—a liquidity effect would not vanish as the threshold widens. That pattern is the threshold-scale confound named in advance in preregistration amendment A3: the night session’s lower volatility makes a fixed five points a coarser sieve, and night $\delta = 5$, $h(1) = 52.88\%$ sits next to day $\delta = 10$, $h(1) = 53.18\%$. What survives is the contrast— $h(1)$ has no trend, $h(k \geq 2)$ declines monotonically. Identifying the cause requires re-running the day–night comparison under volatility-standardised thresholds, which A3 already specifies and this paper does not do.

7 Limitations

- Day session only. The night session exists from 2017-05-16 and is not analysed.
- Brokerage commission is not an official rate; results are reported at NT\$25 per side, with NT\$20 (the exchange floor) and NT\$50 as sensitivities.
- The near-month series is roll-cut, not spread-adjusted, so it cannot support cross-month carry analysis.
- Capacity is measured as size on the fill tick, an upper bound: it assumes the entire print is obtainable and ignores queue priority.
- The preregistration declares `is_blind = false`; the motivating counts were observed before the analysis plan was frozen. Downstream citation must carry that qualification.
- The literature reconciliation covers the scaling-law framework (compared directly on counts and overshoot, Section 4.1) and the price-impact and push-response literature (Section 5.1). It does **not** yet cover the optimal-execution literature, where results 1 and 2 of Section 5.1 could plausibly exist in another form; novelty claims are qualified accordingly.

8 Conclusion

Intraday threshold-crossing predictability in Taiwan index futures is real and, conditioned on incident momentum, large: a 96% reversal rate against an unconditional 51%. It is nonetheless unavailable to a liquidity taker. Execution absorbs 90–107% of the gross edge across two orders of magnitude of threshold, and where it does not, available size does: the median fill when the signal fires is two lots.

The fourteen-year decomposition shows which part of the market changed. The bid-ask bounce component of reversal has no trend, oscillating between 52.58% and 57.92% from year to year, while everything above it declined monotonically and has been competed away. Why the mechanical component stays where it does is not settled here: the day-versus-night test intended to answer it was absorbed by a threshold-scale confound (Section 6, claim 4).

All statistics are computed from the K1824 experiment directory. Every table is reproducible from the committed result JSONs; the entry points are `k1824.py` and `k1824_phase2.py` through `k1824_phase10.py`.

References

- [1] Bouchaud, J.-P., Farmer, J. D., and Lillo, F. (2009). How markets slowly digest changes in supply and demand. *Handbook of Financial Markets: Dynamics and Evolution*, 57–160.
- [2] Glattfelder, J. B., Dupuis, A., and Olsen, R. B. (2011). Patterns in high-frequency FX data: discovery of 12 empirical scaling laws. *Quantitative Finance* 11(4), 599–614.
- [3] Grossman, S. J., and Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *American Economic Review* 70(3), 393–408.
- [4] Guillaume, D. M., Dacorogna, M. M., Davé, R. R., Müller, U. A., Olsen, R. B., and Pictet, O. V. (1997). From the bird’s eye to the microscope: a survey of new

- stylized facts of the intra-daily foreign exchange markets. *Finance and Stochastics* 1(2), 95–129.
- [5] Hasbrouck, J. (1991). Measuring the information content of stock trades. *Journal of Finance* 46(1), 179–207.
- [6] Lo, A. W., and MacKinlay, A. C. (1988). Stock market prices do not follow random walks: evidence from a simple specification test. *Review of Financial Studies* 1(1), 41–66.
- [7] Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *Journal of Finance* 39(4), 1127–1139.
- [8] Vlasiuk, D., and Smirnov, M. (2025). Push-response anomalies in high-frequency S&P 500 price series. arXiv:2511.06177 [q-fin.TR], 9 November 2025.