

Kalshi BTC 15-Minute — The Last 5 Minutes Playbook

Volume 3 — The combined rules: internet research + local mathematics + copper

Direction is the lead · Probability is the Kalshi price · The edge is selection and the fee gate

Built from Volume 1 (internet research, 35 pp) and Volume 2 (local math, 6,482 windows, 19 pp) and the Copper volume (3,656 windows, 20 pp)

Kalshi BTC 15-Minute — The Last 5 Minutes Playbook

Volume 3 of 3. This volume merges Volume 1 (internet research, 35 pages, sourced) and Volume 2 (local mathematics on all 6,482 BTC windows since launch, 19 pages), plus the Copper last-5 volume (3,656 windows, 20 pages), into one set of operating rules.

Every rule is labelled by its evidence:

- **LOCAL** — tested on our own data, chronological 70/30 holdout (models fitted on the first 70% of windows, graded untouched on the last 30%: BTC holdout 1,945 windows, copper holdout 1,097).
- **RULE** — contract arithmetic or a published measurement from Volume 1; true by the contract's terms or another researcher's dataset, not fitted by us.
- **NEXT** — promising in the published evidence, NOT yet tested on our data. Never trade on a NEXT item; it is a test to run, with its falsification criterion stated.

The standing principle of all three volumes: **direction is the lead, probability is the Kalshi price, and the only manufactured edge available is selection — acting only in the tiers that history says hold, and only when the price leaves room after fees.**

1. The BTC rule card (the whole playbook on one page)

1. **Read the lead** — Coinbase 1-minute open at the checkpoint minus the Kalshi strike. Its sign is the call direction. It never gets flipped by any feature in this playbook. **LOCAL**
2. **Take the probability from Kalshi's own %** at that minute (1-minute candlestick Yes open). It was the best-calibrated probability at every minute 5→1 on holdout (Brier 0.130 / 0.115 / 0.094 / 0.074 / 0.034) and no model we built beat it. **LOCAL**
3. **Cross-check with the probit** (Section 3): rolling 60-minute σ , settlement-averaging correction. If the probit and Kalshi % disagree by more than ~10 points, the window is a judgement call, not a tier trade — size down or pass. **LOCAL**
4. **Classify the tier by confidence** (Section 4 tables). $\geq 95\%$ tiers held 97.8–99.3% on holdout; $\geq 85\%$ tiers held 93.4–98.5%. Below 85%, you are in the general population: 81–95% by the clock, dominated by lead size. **LOCAL**
5. **Respect the trap zone:** a lead under \$10 flips 44.9% of the time at 5 minutes left and still 32.2% at 1 minute left. Under \$25, treat the window as undecided regardless of what any single feature says. **LOCAL**
6. **Apply the fee gate before any trade** (Section 5): act only if your probability beats the ask by more than the taker fee at that price plus a buffer. At 99¢ in size the breakeven is 99.07%; a single 99¢ contract rounds its fee to a full cent — 100% breakeven, never worth it. **RULE**
7. **Size with shrunk Kelly** (Section 6): quarter-to-half Kelly on a probability you have already shrunk toward the market price, hard cap 2–5% of bankroll per window. **RULE**
8. **In the final 60 seconds, switch models** (Section 7): the settlement is a 60-second average, so part of it is already locked in. The probit's job there is arithmetic, not prediction. **RULE**

2. What the settlement actually is — and why the last minute is special

RULE (Volume 1, §1, Kalshi contract terms + CF Benchmarks methodology). KXBTC15M settles on the CF Benchmarks Bitcoin Real-Time Index (BRTI), not the daily reference rate. Both numbers that decide the window are 60-second averages of one-second BRTI prints:

- Strike (floor_strike) = average of the 60 BRTI prints in the minute before the window opens.
- Expiration value = average of the 60 BRTI prints in the minute before the window closes.
- Yes wins iff expiration value $\geq$ strike (a tie pays Yes). If data is incomplete, the market resolves No.
- BRTI itself is built from the consolidated **order books** of Bitstamp, Coinbase, Kraken, Gemini, LMAX Digital, Bullish and Crypto.com — Binance, Bybit and OKX are **not** constituents.

Measured consequences (Volume 1, §1.4, DegenHedge on 25,746 windows): a single exchange's spot print points the wrong way on 4.37% of windows, and on 39.2% of windows that finish within \$5 of the strike. The strike equals the previous window's settlement value in 25,281 of 25,285 consecutive pairs — the windows are chained, so the previous settlement is public information the moment a window opens.

Practical meaning for us: our Coinbase lead is a good proxy (it is a BRTI constituent), but near the strike the proxy error is the same size as the lead. That is the deep reason the trap zone in Rule 5 exists, and why the final-minute model in Section 7 works on the average, not the print.

3. The probability model — formulas

3.1 Baseline probit (driftless Brownian bridge)

$P(\text{finishing UP}) = \Phi(\text{lead} / (\sigma \cdot \sqrt{m}))$, where m = minutes left, σ = per-minute \$ volatility, Φ = standard normal CDF.

3.2 The two corrections that help (LOCAL, Volume 2 §8)

- **Per-window σ beats global σ .** Global train $\sigma = \$42.08/\text{min}$, but the rolling pre-window 60-minute σ has median $\approx \$30$ (interquartile $\approx \$20\text{--}41$). Using the rolling σ (or the in-window realized σ) improved probit Brier by 0.009–0.012 at every minute.
- **Settlement averaging.** Because the settlement is an average over the final minute, the variance of (settlement – current price) is $\sigma^2 \cdot (m - 2/3)$ instead of $\sigma^2 \cdot m$: replace $\sqrt{m}$ with $\sqrt{(m - 2/3)}$. At 1 minute left the effective remaining variance falls to one third. This helped the probit at every minute, most at 1 minute (Brier 0.0455 $\rightarrow$ 0.0398).

Even with both corrections the probit (0.135 / 0.120 / 0.100 / 0.082 / 0.040) loses to Kalshi % alone at every minute — which is why it is the cross-check in Rule 3, not the headline number. Its real jobs: (a) sanity-check the market price in illiquid moments; (b) price the final 60 seconds (Section 7), where the market is thinnest and the arithmetic is strongest.

3.3 The same formula in seconds (RULE, Volume 1 §2.4)

For $\tau \geq 60$ seconds left: $\text{Var}(\text{settlement}) = \sigma^2 \cdot (\tau - 40)$, with σ in per-second units — effective time is $\tau - 40$ seconds. This is the identical correction as §3.2 stated in seconds; the internet derivation and our local fit agree.

For $r < 60$ seconds left: a fraction $(60 - r)/60$ of the settlement average is already observed and cannot change. Expected settlement $E[S] = (\text{sum of observed prints})/60 + (r/60) \cdot (\text{current print})$; $\text{SD}(S) = \sigma \cdot r^{1.5}/(60 \cdot \sqrt{3})$. At $r = 30\text{s}$ the SD is $\approx 1.58\sigma$ against a naive $\sqrt{30} \approx 5.48\sigma$ — the outcome is more than three times tighter than a point-settlement model thinks.

3.4 Volatility estimation menu (RULE/NEXT, Volume 1 §2.2–2.3)

- EWMA (RiskMetrics-style) on 1-minute returns — the practical favourite at this horizon; reacts to regime without a window edge.
- Parkinson: $\sigma^2 = 0.361 \cdot (\ln H/L)^2$ per candle; Garman–Klass: $0.5 \cdot (\ln H/L)^2 - 0.386 \cdot (\ln C/O)^2$ — both use the OHLC we already store and are several times more efficient than close-to-close.
- Blend and deseasonalise by time of day; judge any σ change by Brier per minutes-left bucket, not by eye.
- Do NOT use a fixed global σ for live probabilities (see §3.2), and do not use 1-second returns for σ (microstructure noise inflates them — use mids, not trades, where possible).

4. The tiers — what history says each confidence level is worth (LOCAL)

Holdout, best model at each minute. Treat these as upper bounds (the "best model" was selected on the same holdout); the Kalshi-%-alone tiers in Volume 2 §12 are within a few tenths of these everywhere.

4.1 BTC — accuracy on windows at or above the confidence threshold

Minutes left	$\geq 85\%$ conf: accuracy (coverage)	$\geq 90\%$ conf: accuracy (coverage)	$\geq 95\%$ conf: accuracy (coverage)
5	93.4% (52.4%)	95.3% (42.7%)	97.8% (25.2%)
4	95.4% (58.5%)	97.1% (49.4%)	98.5% (34.4%)
3	96.6% (67.4%)	97.5% (59.0%)	98.2% (45.8%)

Minutes left	≥85% conf: accuracy (coverage)	≥90% conf: accuracy (coverage)	≥95% conf: accuracy (coverage)
2	97.4% (72.1%)	98.4% (64.9%)	98.9% (53.2%)
1	98.5% (88.3%)	98.9% (83.9%)	99.3% (78.3%)

4.2 BTC — flip risk by lead size (full sample, 6,482 windows)

Overall flip rate falls 18.7% → 16.4% → 14.4% → 11.5% → 6.0% from 5 to 1 minutes left, but lead size dominates the clock:

Lead	Flip at 5 min	Flip at 1 min
Under \$10	44.9%	32.2%
\$10–25	~30–32%	~15–20% (bucket detail: Vol 2 \$10)
\$25–50	24.5%	3.3%
\$50–75	~14–16%	0.3%
\$75+	under 8%	under 2%
\$150+	~5%	■2% (at 2 min already)

The cliff between 2 and 1 minutes for \$25+ leads is the settlement-averaging signature from §3.2 — it is the single most reliable structural pattern in the local data.

4.3 BTC — signed lead buckets at 5 minutes left (full sample, selected)

Lead bucket	n	Lead-sign accuracy
≤ −\$150	532	95.5–97.2%
−\$150..−100	426	91.8%
−\$100..−75	353	91.2%
−\$75..−50	463	84.4%
−\$50..−25	700	75.6%
−\$25..0	870	52.5–68.3%
0..+\$25	815	58.3–69.9%
+\$25..50	637	75.4%
+\$50..75	439	85.6%
+\$75..100	342	93.3%
+\$100..150	387	95.1%
≥ +\$150	518	95.3–99.7%

Full bucket tables for every minute are in Volume 2 §4.

5. The fee gate (RULE, Volume 1 §4.4)

Kalshi taker fee per contract $\approx 0.07 \cdot P \cdot (1 - P)$, rounded up (fee formula: $\text{roundup}(0.07 \cdot C \cdot P \cdot (1 - P))$ for C contracts).

Breakeven win rate = price + fee:

Entry price	Fee (in size)	Breakeven win rate
90¢	0.63¢	90.63%
95¢	0.33¢	95.33%
99¢	0.07¢	99.07%
99¢, single contract	rounds to 1¢	100% — never trade

Two published post-mortems set the tone (Volume 1 §4.2): at ~3 minutes left, 99¢ asks lost 0.97%, 98¢ lost 2.24%, 90–94¢ lost 8.71% vs 7.81% implied (DegenHedge) — no price group beat its price after fees; and a real-money taker bot grossed +0.67¢ per trade against a 1.55¢ average taker fee. The conclusion is not "don't trade"; it is "the fee gate is the strategy's edge filter": only act when (your probability – ask) > fee + buffer, prefer maker/limit execution where the book allows it, and let the ≥95% tiers (Section 4) do the heavy lifting — a 97.8–99.3% historical tier against a 95–99¢ ask is where the arithmetic can work.

Note the Kalshi-wide favourite–longshot bias (95¢ contracts win ~98% market-wide) does NOT rescue a bad gate: in this market's last 5 minutes specifically, favourites are priced close to fair (above), so the bias is context, not an edge.

6. Sizing (RULE, Volume 1 §6.2)

- Binary Kelly: $f^* = (q - p) / (1 - p)$, with q = your (shrunk) probability, p = price.
- Shrink q toward the market price before sizing (our local result says the market price is the best single probability we have; a model that disagrees strongly is more likely wrong than the market).
- Use quarter-to-half Kelly, hard cap 2–5% of bankroll per window, and skip entirely when the fee gate fails — a skipped window costs nothing; a forced one pays the fee for a coin flip.

7. The final 60 seconds — procedure (RULE, with a build item)

Once $r < 60$ seconds remain, prediction is over; what remains is accounting:

1. Track the running sum of BRTI prints inside the settlement minute. The locked fraction is $(60 - r)/60$ of the final average.
2. $E[S] = \text{locked_sum}/60 + (r/60) \cdot \text{current print}$; $SD(S) = \sigma \cdot r^{1.5}/(60 \cdot \sqrt{3})$ (σ per-second).
3. $P(\text{Yes}) = \Phi((E[S] - \text{strike}) / SD(S))$, with a fat-tail CDF (Section 9, NEXT-2) once tested.
4. Compare against the ask through the fee gate as usual — in the last seconds the market often still prices as if the full minute were live, which is precisely the gap this section exists to measure.

Build item (stated plainly): there is no free real-time BRTI feed (it is a licensed product). The working substitute is a proxy built from the free Coinbase + Kraken order-book websockets (both BRTI constituents), sampled at 1 second. Until that proxy is live, the last fully trustworthy read we produce is the 1-minute mark (the 1:00 tiers in §4.1: $\geq 95\%$ held 99.3% on 79% of windows) — the sub-minute procedure above is specified and ready, and its inputs are the only missing piece.

8. What does NOT help — the negative list (LOCAL + RULE)

Tested on our data (LOCAL, Volume 2) — do not re-add these as direction or probability features:

- Momentum / path (lead change over 1 or 3 minutes): $|\Delta \text{Brier}| \leq 0.0003$.
- Distance travelled, in-window range, pre-window 60-minute return or vol regime as logistic features: ≤ 0.0003 . (In-window vol helps only as σ inside the probit formula.)
- Kalshi % change, Kalshi candle volume / volume change: null to slightly harmful.
- Pooled lead+minutes-left logistic: never beats per-minute lead-only.
- Earlier established nulls at the 6-minute mark: multi-venue lead (Coinbase vs Binance), taker flow/CVD as direction, Polymarket 5-minute cross-signal.

From the published evidence (RULE, Volume 1 §3, §6.3) — no credible 1–5-minute direction value:

- DVOL / 25-delta skew, funding rate level, open interest level, Coinbase premium, stablecoin flows (hours-to-weeks scale, or direction-free; funding/OI are state, not triggers).
- VPIN as a direction signal (direction-free by construction; contested even as a vol predictor — vol gate at best).
- RSI / technical-analysis composites on 1-minute candles as standalone signals.
- Mempool / whale-alert feeds at this horizon.
- Buying cheap longshots or "due" streak bets — the measured streak gap is already ~half priced and smaller than the fee.

9. NEXT — the ranked test queue (promising, untested locally)

Each item has the falsification test that would kill it. None of these is a live rule yet.

- NEXT-1: Order-book imbalance tilt.** Top-5 book imbalance on Coinbase + Kraken (both BRTI constituents): published rank-IC +0.113 at 1-minute forward, extreme deciles $\pm 1.1\text{bp}$ on the next minute — inside the taker fee as a trade, but viable as a bounded tilt on the probit/logit probability. Test: log live books (the book logger already exists), add OBI as a capped logit adjustment, require holdout Brier improvement ≥ 0.002 at 3–1 minutes without accuracy loss.
- NEXT-2: Fat-tail CDF.** Replace Φ in the probit with a Student-t ($v \approx 3-5$) or jump-mixture CDF. Motivation: 1-minute BTC returns are severely fat-tailed, and favourites at 90–99¢ are exactly where the tail lives. Test: same probit, swap only the CDF; require Brier improvement concentrated in the $\geq 90\%$ buckets.
- NEXT-3: σ blend.** EWMA + Parkinson/Garman–Klass blend, time-of-day deseasonalised, HAR-style short/long RV ratio. Test: drop into §3.2's probit in place of rolling σ ; require Brier gain in every minutes-left bucket, not just on average.
- NEXT-4: The BRTI proxy + final-minute accounting.** Build the 1-second Coinbase+Kraken proxy, reconstruct historical settlement minutes, and backtest Section 7's procedure directly. Test: on windows finishing within \$25 of the strike, the sub-minute model's Brier must beat the 1:00 Kalshi % — that subset is where it can possibly matter.
- NEXT-5: Maker-first execution study.** Measure achievable fills posting at the model price vs taking, on logged books. The measured gross leader edge (+0.67¢) is smaller than the taker fee (1.55¢), so execution style may be worth more than any remaining signal.

10. Copper — its own rules (LOCAL, Copper volume)

Copper (KXCOPPER15M, settles on Pyth; our spot is a Binance/Yahoo proxy) shares the skeleton but reverses two BTC rules. Do not port BTC rules to copper without this section.

- On conflict, trust Kalshi % over the move.** Kalshi's % sign beat the (self-consistent) move sign at every minute on holdout: 79.9 / 83.5 / 85.2 / 88.5 / 92.0% vs 78.6 / 79.8 / 81.8 / 83.0 / 86.0%. When they disagree, Kalshi is right 55.2% at 5 minutes, rising to 75.0% at 1 minute. The proxy-vs-Pyth basis is the reason, and it does not shrink into the close.
- No settlement-averaging correction.** It hurt copper at 1 minute (probit Brier 0.0949 $\rightarrow$ 0.1026) for every σ choice — sharpening around a proxy move is overconfidence. Per-window σ was also null; global $\sigma = 0.00203/\text{min}$ is the best of a weak probit family (mean Brier 0.121 vs Kalshi 0.100).
- A move under 0.002 is no lead.** It flips 36.3% at 5 minutes and still 32.6% at 1 minute; sub-0.002 moves hold only 59–64% at every minute vs 95–98% for moves ≥ 0.005 . There is no BTC-style 2 $\rightarrow$ 1 minute flip collapse.
- Best model** = logistic(move + Kalshi %), but its edge over Kalshi % alone is 0.0010–0.0018 Brier — below the "helps" bar: consistent, marginal. Everything else null, same list as BTC §8.
- Copper tiers (Kalshi % alone, holdout):**

Minutes left	$\geq 85\%$: acc (cov)	$\geq 90\%$: acc (cov)	$\geq 95\%$: acc (cov)
5	94.8% (47.0%)	95.3% (37.2%)	97.8% (24.5%)
4	94.4% (54.1%)	95.7% (44.8%)	97.3% (33.5%)
3	95.4% (60.9%)	96.6% (53.9%)	98.2% (41.2%)
2	95.7% (70.6%)	96.7% (62.7%)	98.2% (51.8%)
1	97.5% (80.4%)	98.2% (74.9%)	98.7% (67.9%)

- Copper sample caution: 3,656 windows (weekdays only, since Aug 27) — every copper number carries roughly $\sqrt{2}$ the uncertainty of the BTC numbers.

11. The live checklist (per window, last 5 minutes)

At 5:00 left, and refreshed each minute:

- Lead \$ and sign (Coinbase open vs strike) — direction.
- Kalshi % — the probability. Note the tier (\$4.1) and the lead bucket (\$4.3).
- Probit cross-check (rolling σ , $\sqrt{(m - 2/3)}$) — flag if it disagrees with Kalshi % by >10 points.
- Trap check: $|\text{lead}| < \$10 \rightarrow$ undecided; \$10–25 $\rightarrow$ low tier regardless of features.

- If trading: fee gate (§5) → shrunk-Kelly size (§6) → limit/maker first.
- Under 60 seconds: Section 7 accounting (once the BRTI proxy is live); until then, the 1:00 read is the final word.
- Copper windows: Section 10 overrides — Kalshi % wins conflicts, no settlement-averaging, sub-0.002 = no lead.

12. Open items and honest discrepancies

- **Series start date.** Our data begins Aug 13, 2026 and direct Kalshi series queries return zero KXBTC15M markets before that; Volume 1 found public records suggesting similar markets from Dec 2025/Jan 2026 (the 60-second-average rule effective Jan 10, 2026). Unresolved; every local number in these volumes is computed on the Aug 13 → present data we can actually pull, and says so.
- **Polymarket fees.** Volume 1 found the current fee formula conflicting across sources by date; do not hard-code a Polymarket fee from these volumes.
- **BRTI access.** No free real-time BRTI API exists (licensed). Section 7's proxy is the substitute; its accuracy vs true BRTI is itself a NEXT-4 measurement.
- **Selection caveat.** Tier accuracies in §4 are historical holdout rates for the stated subsets, selected as described in Volume 2 §11 — they describe those tiers, not a promise about any single window.