Team 12 — QuantiFi 2026

Team Members: Sunny Zhang (Team Lead), Manas Kohli, Katherine Shaw, Bosco Yu

Overview

We trade all 26 products with a single execution engine and a separate fair value (FV) model for each product family. On the three training days the trader makes 162.1k, 160.1k and 166.6k (488.8k in total) in the official engine, and 477.9k under the stricter `worse` fill rule. The website backtester gives the same result (488,750). Every one of the 30 blocks of 1,000 ticks is profitable, and the maximum drawdown of total PnL is 11.7k. Because all parameters were chosen on these same three days, the last section describes how we tested for overfitting.

Execution

The simulator does not react to our orders, and for 25 of 26 products every trade prints exactly at the best bid or ask. Our first trader, a market maker quoting one tick inside the spread, lost money on most days: pennying won no additional fills, and in TELLURIUM and the vouchers, which account for about 90% of traded volume, adverse inventory cost more than the spread earned. Joining the touch instead of improving it was worth about 10k a day. Since queue position carries no value here, the edge has to come from the fair value.

Each tick, for each product, the engine takes any quote mispriced against FV by more than a threshold, clears inventory at FV, joins the best bid and ask (capped at FV $\pm$ edge), and skews quotes by $-kq$ for position q. Order sizes are checked against the position limit separately for each side, so the exchange never cancels a tick's orders. Each product runs in its own `try/except`, and all state is carried in `traderData` ($\approx$1 KB).

			PnL by day (k)			
Family	**Products**	**Fair value**	-2	-1	0	**Total**
Jump regime	TELLURIUM	Mid EMA; after a jump, lean ± 30 in the confirmed direction	53.2	47.8	67.2	168.1
Options	10 vouchers, LYCHEE	Black-Scholes at own IV and at the smile IV, plus order flow	53.4	48.3	39.2	140.9
Anchors	AMETHYST, TOFFEE_CAKE	Fixed at 10,000 / 5,000	21.6	24.3	23.2	69.1
Trend	PAPER_LANTERNS, GOLDEN_PEARLS	Fast EMA $+\,\beta\,$(fast $-$ slow EMA)	20.3	21.9	21.8	64.0
Pearl identity	PINK, WHITE, BLACK	$50{,}000 - \Sigma$ other pearls; lean during GOLDEN steps	10.3	11.0	11.1	32.5
Breakout	BLUE_PEARLS	Cross to $\pm$limit when $\lvert\text{mid} - \text{EMA}\rvert > 7$	4.5	5.2	5.1	14.8
Breadth	6 thin products	Mid, small size, tight inventory	-1.2	1.4	-0.9	-0.7
Total			**162.1**	**160.1**	**166.6**	**488.8**

TELLURIUM

TELLURIUM is the largest pool. Our first model, a trend-following FV, was unstable: changing one parameter from 1.0 to 0.5 moved a single day's PnL from +29k to $-$33k. The cause is jumps. Eight to thirteen times a day the mid moves 77–300 in one tick, against at most 14 on any other tick, and then drifts about 90 further over roughly 200 ticks. The drift continues the jump in two cases out of three and reverses it in the third, so the jump itself carries little directional information.

The drift's direction becomes clear early. Once the mid has moved 12 points from its post-jump level (about two standard deviations of a normal tick), the direction of that move matches the rest of the drift in 27 of 30 training jumps (Figure 1). An earlier version confirmed direction after a fixed number of ticks instead; tuned on two days, it made only 24k on the held-out third, and we discarded it.

The trader detects a jump as $|\Delta m| > 40$, keeps quoting around an EMA of the mid, and waits up to 25 ticks for a 12-point move. If one occurs, FV becomes $\text{EMA}(m) + 30d$ for 400 ticks, where $d = \pm 1$ is the confirmed direction, and the engine builds the position to the limit. TELLURIUM PnL rose from 98k to 168k and improved on every day. Any jump threshold from 25 to 70 gives identical results.

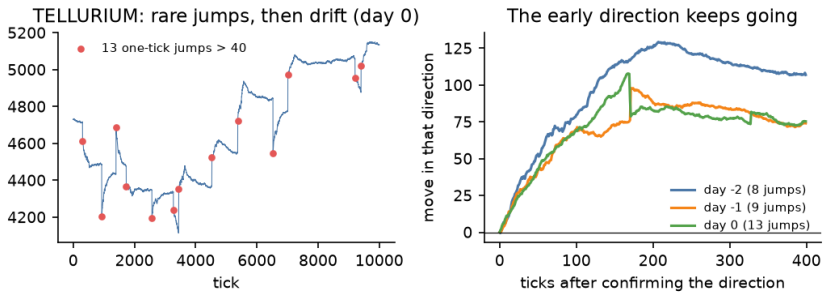


Figure 1: Left: day 0, with the one-tick jumps marked. Right: the average path once the post-jump direction is confirmed, signed by that direction, for each day.

Options

Quoting the ten vouchers around a Black-Scholes value lost 30k over the three days. The average fill moved from $+1.5$ to -2 over the next 100 ticks even after delta-hedging with LYCHEE, so the flow trading against us was informed about the vouchers themselves. A full hedge was not available in any case, since 500 voucher contracts far exceed LYCHEE's limit of 200, and we found nothing that predicts LYCHEE.

Two variables do predict a voucher's next move. The first is its own recent signed order flow, which correlates with the next 50-tick move at $r = 0.2$–0.4 on strikes 7300–8100.

The second is its distance from a volatility smile fitted across all strikes each tick,

$$\sigma(x) = a + bx + cx^2, \qquad x = \frac{\ln(K/S)}{\sqrt{T}}.$$

On strikes 7200–7900, this distance predicts the delta-hedged 100-tick move at $r = -0.35$ to -0.53, against -0.15 to -0.26 for the distance from each strike's own rolling implied volatility, on every day (Figure 2). The furthest out-of-the-money strike, 8100, is the exception. Adding a learned per-strike offset from the smile reduced PnL, so it is not used.

The voucher fair value is

$$\text{FV} = m + 0.25\,\big(\text{BS}(\sigma_{\text{own}}) - m\big) + 0.4\,\big(\text{BS}(\sigma(x)) - m\big) + \gamma_K F,$$

where m is the mid, σ_{own} is an EMA of the strike's implied volatility and F is signed flow with a 20-tick memory. Net delta across all vouchers and LYCHEE is kept within ± 50. Options PnL went from -30k to 131k with the flow term and to 141k with the smile.

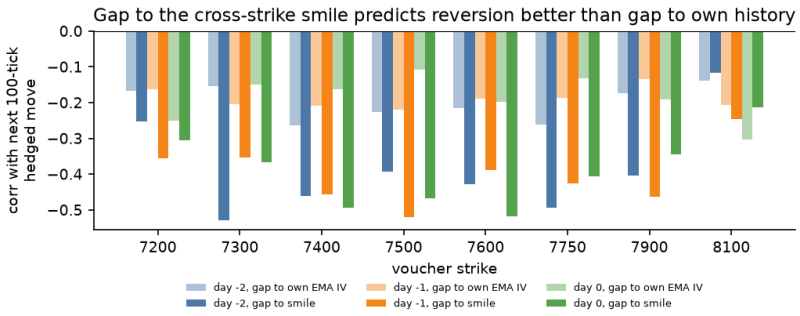


Figure 2: Correlation of each strike's mispricing with its next 100-tick delta-hedged move. Dark bars: distance from the cross-strike smile. Light bars: distance from the strike's own rolling IV. More negative means stronger reversion.

Pearls

Individually the five pearls look like random walks, but their mids sum to 50,000 (± 5) on every day, and PINK is the leg that closes deviations. PINK's fair value is therefore $\frac{1}{2} m_{\text{PINK}} + \frac{1}{2}(50{,}000 - \sum_{\text{others}} m)$.

The identity also holds during GOLDEN's three level steps, when GOLDEN falls about 2 per tick for roughly 45 ticks and WHITE and BLACK rise to offset it. A smoothed GOLDEN velocity identifies each step within 1–3 ticks and never triggered outside the three real steps; during a step we shift the WHITE and BLACK fair values in the offsetting direction. This doubled their PnL from 2.2k to 4.4k, and the three pearls make 32.5k in total.

Anchors and trends

AMETHYST (standard deviation 0.4 around 10,000) and TOFFEE_CAKE (daily means 5,004, 5,003 and 5,002) are traded around fixed values. If inventory stays at the limit for 2,000 ticks, the FV falls back to a slow EMA, which protects against a shifted anchor. TOFFEE's original take threshold of 5–10 lay where PnL was highly sensitive to the parameter; 12 lies in a stable region and is better on every day.

PAPER_LANTERNS has lag-1 autocorrelation of -0.5, but this reflects about ± 1 of noise on a piecewise-linear trend: changes over consecutive 100-tick windows correlate at 0.78, 0.81 and 0.82 on the three days. Its FV extrapolates the gap between two EMAs, $\text{FV} = E_f + \beta(E_f - E_s)$, with smoothing factors 0.1 and 0.02 and $\beta = 4$. GOLDEN_PEARLS uses the same model. It could not be anchored at 8,000, because it stepped to 7,900 and then 7,750 and each day opens where the previous one closed.

BLUE_PEARLS

BLUE's 10-tick returns correlate at about 0.23 on every day, yet shifting quotes in proportion to recent momentum made between -1.7k and $+0.7$k. BLUE prints only about 60 trades a day, which is too few to build a position passively, and the fills we did get were adverse. Signed flow had no predictive value, and BLUE has no jumps (its largest one-tick move is 7.5).

The momentum is nonlinear (Figure 3). With $s = m - \text{EMA}_{0.2}(m)$, deviations below 4 show no follow-through, while deviations above 6 (about 3 standard deviations) continue another 12–20 points, against a spread of 6. BLUE therefore posts no quotes. When $|s| > 7$ the trader crosses the spread to the position limit in the direction of s, and it exits when s changes sign. This earns 4.5–5.2k a day, up from roughly zero. The same test on every other product found no pattern that held on all three days.

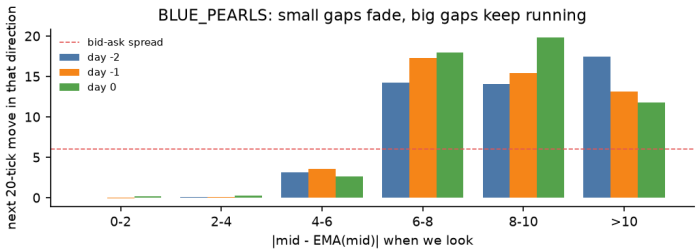


Figure 3: Mean move over the next 20 ticks, in the direction of the deviation, by deviation size. Only deviations larger than the spread (dashed) are profitable to trade.

Rejected ideas

Basket against its components. FRUIT_BASKET's premium over its three fruits moved from 6,800 to 7,700 to 7,650 across the days, and the best-fit weights changed each day, so we did not trade the basket against its components. With 15–17 basket trades a day, it is traded in small size for breadth.

Predicting LYCHEE. LYCHEE showed no predictable component, so it is quoted around its mid and used to manage voucher delta.

Overfitting

Every parameter was chosen on the same three days we report, so the backtest overstates what to expect on the scoring day. Most of our 100-tick statistics also use overlapping windows, which makes the effective sample far smaller than the tick count. We therefore kept a change only if it passed four tests.

1. Consistency across days. The effect had to appear on each day separately. The TELLURIUM drift, the smile residual, the pearl identity and the BLUE tails all do. The basket relationship and the breakout rule on other products did not, and were left out.

2. Parameter stability. The chosen parameters had to lie inside a stable region. Every BLUE setting adjacent to ours makes 11–16k and is positive on every day, and the TELLURIUM grid stays between 163k and 172k. For BLUE and TOFFEE the in-sample optimum was at the edge of the stable region, and we chose its centre at a small cost in backtest PnL. Only the pearl sum and the AMETHYST and TOFFEE levels are hard-coded; all other levels are estimated online.

3. Leave-one-day-out validation. A model tuned on two days had to hold up on the third. This rejected the time-based TELLURIUM design, which made only 24k on the held-out day. An earlier version of the full trader scored 115.1k on held-out day 0 against 116.9k in-sample.

4. Stress tests. Results had to survive misspecification. With the TOFFEE anchor wrong by 30, TOFFEE still makes 24–27k. Assuming voucher expiry anywhere from 3 to 8 days changes options PnL by less than 1%. The stricter `worse` fill rule costs 2%.

The two regime models hold at most the position limit and enter only after a move has begun, so if a pattern fails on the scoring day the cost is about one spread per event, roughly 4k a day for TELLURIUM. We expect the scoring day to earn less than the backtest, but these checks suggest the shortfall should be modest.