PREDICTION-MARKET MICROSTRUCTURE · CASE STUDY · 24 JULY 2026

Analysis of Polymarket France–Spain and World-Cup-Winner markets at match time

Maxime Scali

Abstract. We study the joint behaviour of two families of Polymarket contracts around the FIFA World Cup semi-final between France and Spain (14 July 2026): the markets written on the match itself and the world-cup-winner market. Using a complete tick-level trade record reconstructed from the Polygon blockchain, we measure how goal-driven repricing in the match market transmits to the winner market. Transmission is fast, half-complete within 10–20 seconds, and of the theoretically correct magnitude. The clearest deviation we document is a delayed re-rating of the losing team's conditional strength, roughly 12c over the hour after the first goal, with no news event to trigger it; the non-playing teams' contracts, by contrast, move only a few cents as the four-team sum rebalances around one. Both books begin trading almost instantly after each goal, within the first two-second block.

Motivation. We choose this pair of markets, the France–Spain match book and the world-cup-winner book, for three reasons.

- **Volume.** The markets on this single match cleared over $130M, and the world-cup-winner market has traded roughly $4.3B across its team legs since inception, so no mispricing documented here can be blamed on a thin venue.
- **Identification.** A knockout tie and the winner market are linked by a no-arbitrage relation that involves no model: the conditional probability linking the two books is directly computable from traded prices (equation (2)), so any deviation is a genuine cross-market inconsistency rather than a modelling artifact.
- **Measurability.** Goals are sharp, precisely-timed public shocks, which lets transmission speed and magnitude be read to the second.

Setting. The semi-final was played on 14 July 2026, kickoff 19:00 UTC; Spain won 0–2, with goals at 19:19:30 and 20:19:30 UTC and the final whistle at 20:49 UTC. The match book refers to the contracts settled by the evening's result, principally the team-to-advance market: one share of the winning side pays $1 when its team reaches the final. The winner book refers to the world-cup-winner market: one contract per team, paying $1 if that team wins the tournament, of which four legs were alive at kickoff: Spain, France, England and Argentina. Trading is not costless. Fees fall only on the taker (the party crossing the spread); resting maker orders pay essentially nothing. The taker fee follows $0.05 \times n \times p \times (1-p)$ for n shares at price p, peaking at $1.25 per 100 shares when $p = 0.5$. The team-to-advance market alone generated roughly $0.48M in fees.

Data. Three sources, collected for this study: (i) a complete tick-level trade record for 11–15 July 2026: 846,502 fills reconstructed from the OrderFilled events on Polymarket's Polygon exchange contracts, each with a block timestamp (≈2s), price, size, and wallet; the blockchain is the only complete public source, and it matches the API exactly wherever the API is complete. (ii) Official one-minute closing prices, for horizons beyond the match. (iii) Contract metadata. Prices are read as probabilities and quoted in cents (1c = $0.01).

Contents

- The no-arbitrage relation between the two books and the pre-match state

1 Do the theoretical relations between the books hold?, the France+Spain sum against its null, and the consistency of relative odds

2 The non-playing teams, how England and Argentina moved during and after the match

3 Timing and intensity: transmission speed at the goals, activity in the first seconds, and a concluding synthesis

The no-arbitrage relation between the two books and the pre-match state

Exactly one of the two teams advances from a knockout tie. Let p_S denote the probability that Spain advances, as priced by the match book, and $p_F = 1 - p_S$ the corresponding probability for France. For each team $i \in \{F, S\}$, let W_i denote the probability that team i wins the World Cup, as priced by the winner book, and let q_i denote the probability that team i wins the final conditional on reaching it. Winning the cup means advancing and then winning the final, so the two books are linked by

$$W_i = p_i\, q_i\,, \qquad S = W_F + W_S = p_F\, q_F + p_S\, q_S \quad (1)$$

Both books trade continuously, so the conditional strength is itself observable at every instant as a ratio of traded prices,

$$q_i(t) = W_i(t) \,/\, p_i(t) \quad (2)$$

with no model or estimation involved. Over the pre-match afternoon (medians over 12:55–18:50 UTC) these ratios stood at $q_F^0 = 0.663$ and $q_S^0 = 0.529$: the market regarded France as the clearly stronger prospective finalist.

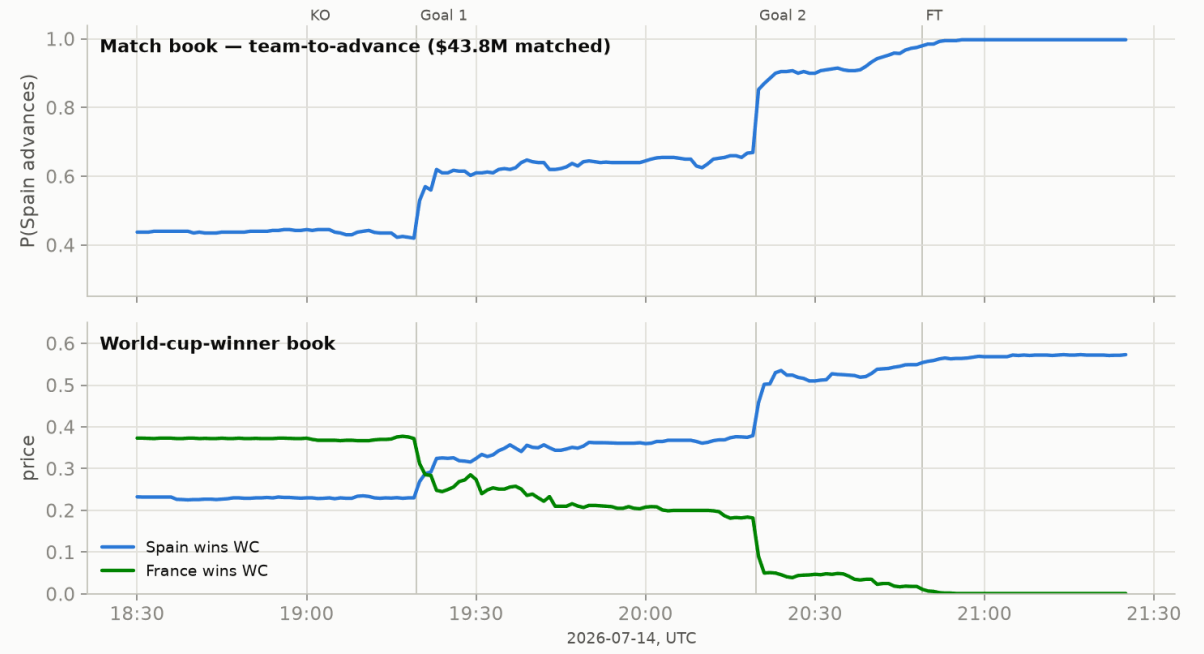


Figure 1. The same evening in the two books, on the synchronized 60-second same-bucket median. Top: $p_S(t)$ (16,435 trades). Bottom: the winner-book legs $W_S(t)$ and $W_F(t)$ (4,222 and 8,748 trades). Vertical lines mark kickoff, the two goals and full time.

Part 1 · Do the theoretical relations between the books hold?

Consistency of relative odds across the books

Here we test whether the two books price consistent relative odds. If they price the same world, the ratios $q_i(t)$ of equation (2) should move only on news about the final; ratio levels are read from the synchronized 60-second same-bucket median of both books. This means we slice time into one-minute buckets and, for each book, take the median trade price inside each bucket, so the two legs of the ratio are measured over the same window rather than at their two most recent (and generally non-simultaneous) trades. Full time gives a clean check: with $p_S = 1$, "Spain advances then wins the final" and "Spain wins the cup" are the same claim, so q_S and W_S must coincide.

They do, to a tenth of a cent ($q_S = 0.5734$ against $W_S = 0.572$), and q_S otherwise only steps up on the goals, re-rating on evidence rather than inconsistency. France's leg (Figure 2) separates into two effects. At the goal the books repriced ratio-preservingly: synchronized q_F is essentially unchanged (0.651 to 0.657), and the post-goal excursion toward 0.74 in the last-trade ratio is likely a measurement artifact, since the more granular non-median series oscillates heavily around the 0.67 level rather than settling higher. A genuine, unexplained variation remains in the level of q_F, which declined about 12c over the hour after

the goal, with no goal between, as France's winner leg drifted down while the match book barely moved. Two readings fit and one match cannot separate them: gradual incorporation of soft information (an hour of poor play is itself evidence about a hypothetical final) versus a sticky longshot leg sold down slowly.

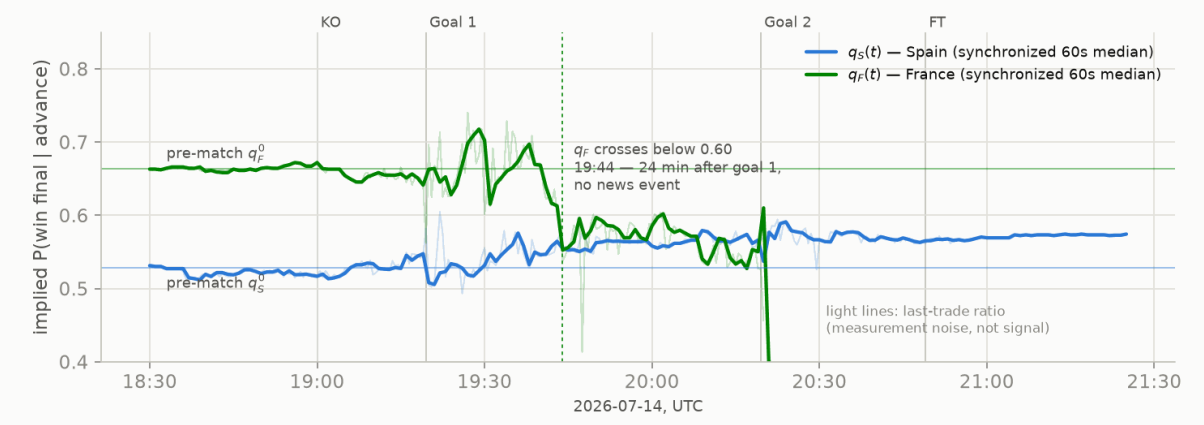


Figure 2. The traded conditional strengths $q_i(t)$. Bold lines are the synchronized ratio, the same-bucket median of both books at 60-second resolution, from which the levels in the text are read; the faint lines behind them are the 30-second last-trade ratio, whose heavy post-goal oscillation is measurement noise rather than genuine disagreement. Pre-match levels are shown as reference lines; the dashed marker is the durable crossing of q_F below 0.60 at 19:44, 24 minutes after goal 1. The France series is dropped once p_F becomes tiny, where ratio (2) divides two small noisy numbers.

> When the market is calm the two books agree to a tenth of a cent, and at the shock itself relative odds were preserved, so the apparent post-goal disagreement is measurement noise. The one unexplained deviation is the losing team's conditional strength, which re-rated about 12c over the hour after the goal with no news event to account for it.

The France + Spain sum against its null path

Here we test whether the sum S moved more than a normal day warrants, and whether that movement is justified. As a scale we use total variation, the sum of absolute one-minute changes, in cents per hour. To judge the movement we compare S with the null path

$$\hat{S}(t) \;=\; p_F(t)\, q_F^0 \;+\; p_S(t)\, q_S^0 \qquad (3)$$

the value the sum would take if both conditional strengths had stayed at their pre-match levels and only the advance probability had moved. During the match total variation reached 28 c/h against a 3.1 c/h baseline, a factor of nine, yet S tracks the null closely (Figure 3) and ends above it, falling only to 0.573, not toward $q_S^0 = 0.529$, because traders re-rated Spain's conditional strength upward as it dominated (q_S rose 0.529 to 0.572). The remaining departures are the

minute-scale goal-time spikes, mechanical artifacts of the two books never repricing at the same instant.

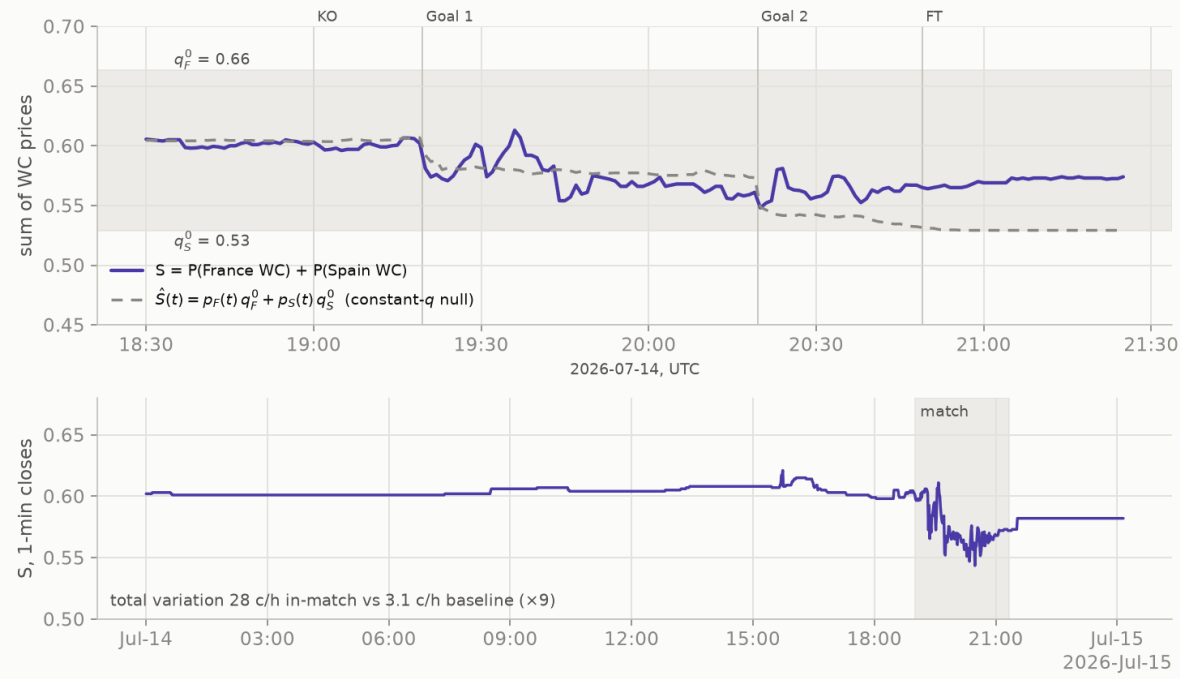


Figure 3. Top: S on the synchronized 60-second same-bucket median against the null path $\hat{S}$ of equation (3); the shaded band is $[q_S^0,\ q_F^0]$. Bottom: S in one-minute closes across 14 July, nearly flat outside the match (shaded).

> The sum moved nine times more than on a normal day, but almost all of it is the drift predicted by (1) plus an upward re-rating of Spain; the mispricing component reduces to minute-scale spikes at the goals.

Part 2 · The non-playing teams

Here we test whether the legs of the teams not playing moved as the identity implies. Two channels justify some movement: Spain advancing means the survivor faces the weaker prospective finalist (since $q_F^0 > q_S^0$), good news for both; and because the four winner legs sum to roughly one, the drop in $W_F + W_S$ must be absorbed by England and Argentina between them. Both channels predict modest, positive moves concentrated at the goals.

The observed moves match this picture (Figure 4). Both legs are more active during the match than in the quiet pre-match hours, but the amounts are small: a few cents at most, consistent with the two legs simply absorbing the change in $W_F + W_S$ so that the four-leg sum stays near one. Over the longer term the net effect is modest and positive for both: comparing the pre-match level with the day after the match, England gains roughly a cent and Argentina about two.

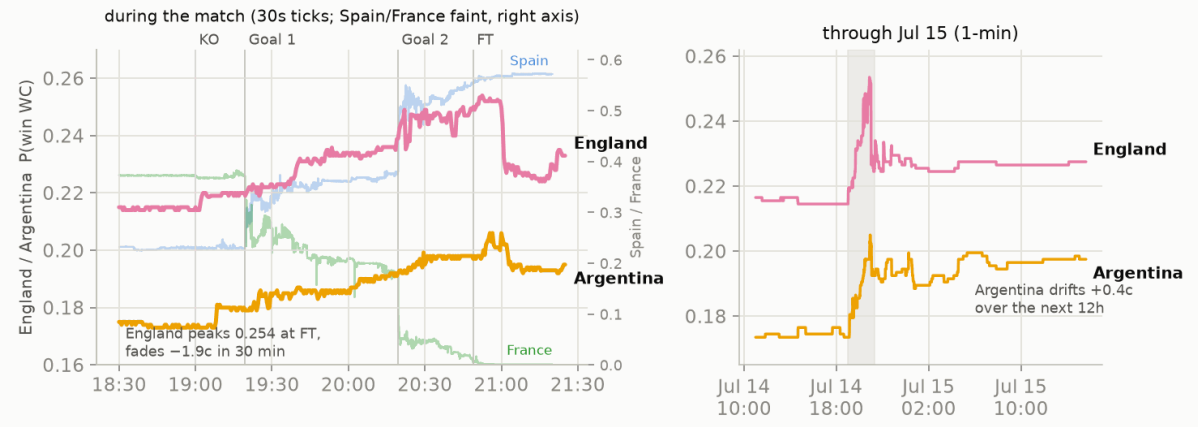


Figure 4. Left: all four winner legs during the match (30-second ticks), with England and Argentina in the foreground on the left axis and the two playing teams, Spain (advancing) and France (eliminated), shown faintly on the right axis for context. Right: England and Argentina through 15 July (one-minute closes; the match window is shaded), showing the small net gain each leg holds after the match.

> The non-playing legs move mainly to keep the four-team sum near one. That movement is larger during the match than before it, but small in absolute terms, and the net before-to-after effect is a gain of a few cents for both England and Argentina.

Part 3 · Timing and intensity of transmission

Size and speed of transmission at the goals

Identity (1) makes a quantitative prediction: a goal that moves the advance probability by Δ should move Spain's winner price by $q_S\Delta$ and France's by $-q_F\Delta$. We define an event as a 60-second median move of the advance book exceeding 6c; the two goals are the only such events, each anchored to its true clock time (19:19:30 and 20:19:30 UTC).

For each book we build one 5-second price series around each goal; the table below and the response curves of Figure 5 are both read from it. Every trade is a taker-side on-chain fill, priced as USDC paid per token and time-stamped by its Polygon block (≈2-second resolution). We lay a grid of points spaced 5 seconds apart from 300 s before the goal to 300 s after it,

$$t_k \;=\; t_{\text{goal}} + 5k \text{ s}, \qquad k = -60, \ldots, 60$$

and set each grid point to the median price of the trades falling in the preceding 5-second bucket, carrying the last value forward across empty buckets:

$$p(t_k) \;=\; \text{median}\{\, \text{fill prices in } (t_k - 5\,\text{s},\, t_k] \,\}$$

The per-bucket median suppresses the bid–ask bounce, the small oscillation between buyer and seller-initiated fills, which the normalization below would otherwise magnify into several points of spurious wiggle. A pre-goal baseline p_0 is the median over roughly the four minutes before the goal and a settled level p_1 the median over a four-minute window centred five minutes after it. The table reports the total repricing $p_1 - p_0$ in cents; the response curve is the same series normalized so that 100% is the eventual move,

$$r(t) = \frac{p(t) - p_0}{p_1 - p_0}$$

and the delays t_{50} and t_{90} are the first times $r(t)$ crosses 0.5 and 0.9.

BOOK	GOAL 1 MOVE	GOAL 1 T_{50} / T_{90}	GOAL 2 MOVE	GOAL 2 T_{50} / T_{90}	MAGNITUDE CHECK
● match (advance)	+19.2c	5 s / 165 s	+24.0c	5 s / 50 s	reference
● winner, Spain	+9.5c	20 s / 130 s	+14.8c	10 s / 125 s	9.5/19.2 = 0.50 and 14.8/24.0 = 0.62, both $\approx q_S$ ✓
● winner, France	−12.4c	10 s / 130 s	−13.8c	10 s / 30 s	87–97% of $q_F^0\Delta$; consistent with an unchanged ratio at the goal, q_F's later decline is the subject below

The magnitudes are the predicted ones, and the delays are measured in seconds, generally shorter at the second goal. The correlation of ten-second returns confirms the ordering of the books: 0.35 contemporaneously, 0.13 with the match book leading by one step, and approximately zero in the reverse direction and in quiet play, so the books re-align only when something happens. Figure 5 shows the full response curves.

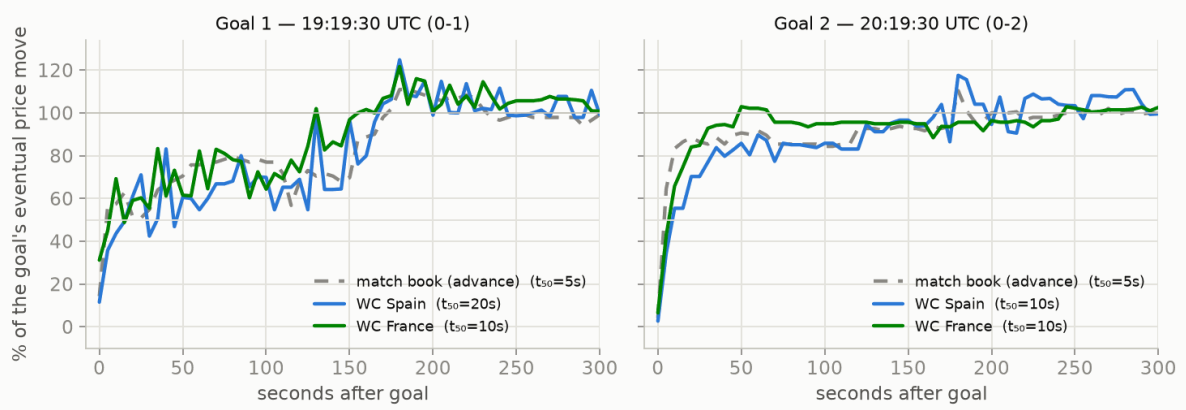


Figure 5. Percentage of the goal's eventual price move completed as a function of time since each goal. Each series is the per-5-second-bucket median price, normalised by the total move settled over the following five minutes; a point at 50% means the book had travelled half its eventual distance. The winner-book legs track the match book (dashed) within seconds; France's leg overshoots (above 100%) at goal 1 before settling.

> Goals reached the winner book at the size implied by (1), half-complete within 10–20 seconds and 90% complete within about two minutes; transmission is fast, event-driven, and quantitatively correct.

Trading intensity in the first seconds

The response curves smooth over the very first seconds, which the block-level record resolves (Figure 6). Activity in all three books sits near zero until the goal, then jumps in the first block or two at once: the advance book carries about 137 and 115 fills in the first ten seconds of the two goals, against a handful beforehand. The two goals differ in character, goal 2 concentrating into a single $\approx$ \$57k advance-book block while goal 1 is spread over several. Goal 1 was a penalty, whose drawn-out sequence (award, wait, run-up and kick) staggers the information over several blocks, which helps explain both its more spread-out volume and the messier price discovery of Figure 2, against the single sharp shock of goal 2.

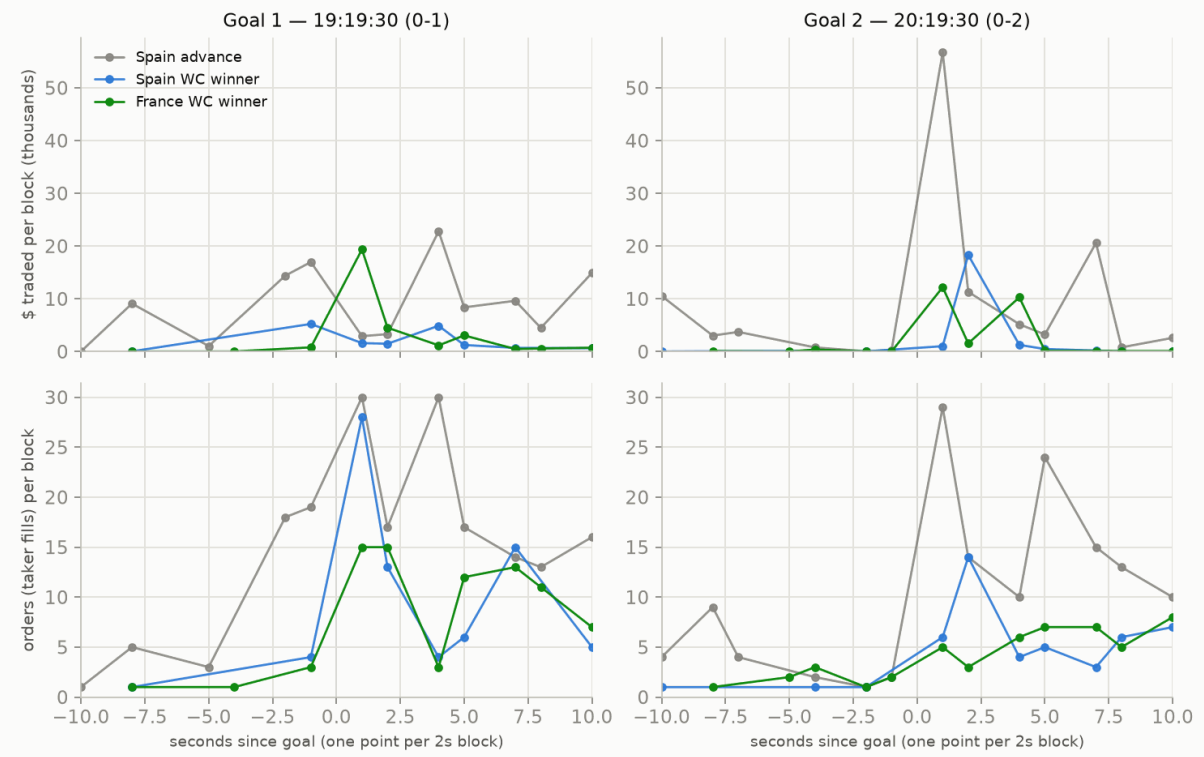


Figure 6. Trading intensity in the twenty seconds around each goal, one point per 2-second block. Columns are the two goals; the top row is dollar volume traded per block (thousands), the bottom row the number of orders per block; the three books are the advance book and the Spain and France winner legs. An order here is a taker execution: one row per market order matched, with the resting maker legs excluded to avoid double-

counting, and the ≈2-second Polygon block is the finest available timestamp: sub-block timing is not public, though execution order within a block is exact.

Conclusion

Across a single match the two books stayed largely consistent over time: the identity held before kickoff, the sum tracked its null path, and goals transmitted to the winner book at the size and speed the identity predicts, half within ten to twenty seconds. The apparent goal-time dislocations dissolve on synchronized measurement, being artifacts of two books never repricing at the same instant, and the small residual inconsistencies were in any case too small to capture net of the taker fees. The one genuine inaccuracy is the losing team's conditional strength, which decayed about 12c over the hour after the first goal with no news to account for it. All of this rests on one match; the mechanisms are the finding, and the effect sizes await replication on the remaining fixtures.

An interesting follow up would be to apply the same analysis to markets whose conditional structure is far less obvious than a knockout tie, such as geopolitical markets, where the relation linking related contracts is not given by a clean identity and would itself have to be inferred.