© JUN 2026 | IRE Journals | Volume 9 Issue 12 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I12-1718842

# Exploiting Wicket-Driven Mispricings in Cricket Prediction Markets

AVIRAL SHARMA[1], CHANDRIMA SABHARWAL[2]
[1]*Meta*
[2]*Amazon*

***Abstract- Prediction markets are exchange-traded contracts that pay out based on the outcome of a future event, with prices reflecting the market's consensus probability of that outcome. When new information arrives, prices should adjust instantaneously under the efficient market hypothesis. In practice, however, thin limit-order-book markets may reprice gradually, creating transient windows in which informed participants can trade at stale prices. We study this phenomenon in the context of live Twenty20 (T20) cricket, where a wicket—the dismissal of a batter—is an instantaneous, unambiguous information event that discontinuously revises match-win probabilities. We develop a multiplicative wicket-impact signal model, calibrated on 353 Indian Premier League (IPL) matches and over 4,300 wicket events, that estimates the magnitude of the expected price shift as a function of player quality, match phase, innings context, and recent form. The model is paired with a systematic trading strategy that enters positions on the fielding team's contract immediately after high-impact dismissals, exploiting the delay between the information event and full market repricing. We evaluate the strategy through three progressively demanding tests: a historical backtest across 52 IPL 2025 markets, an out-of-sample validation on 17 ICC Men's T20 World Cup 2026 markets, and a live deployment spanning the full 64-match IPL 2026 season. Across all three settings, the strategy produces statistically significant positive returns. An innings-level decomposition further reveals that the repricing inefficiency is concentrated in the second innings, where wickets have a more direct and calculable effect on the probability of reaching a known target. These findings provide evidence that prediction markets for live sporting events exhibit systematic, exploitable inefficiencies in the immediate aftermath of discrete information shocks.***

## I. INTRODUCTION

The efficient market hypothesis [Fama, 1970] posits that asset prices fully and instantaneously reflect all available information. In prediction markets—where contracts pay $1 if a specified event occurs and $0 otherwise—this implies that prices equal the market's best estimate of the event probability, updated immediately upon new information. While considerable evidence supports aggregate calibration in prediction markets [Wolfers & Zitzewitz, 2004, Arrow et al., 2008], efficiency at the aggregate level does not preclude localised, transient deviations when information arrives in discrete and sudden bursts.

Live cricket presents precisely such an environment. A wicket—the dismissal of a batter—is an instantaneous, unambiguous event that discontinuously revises the conditional win probability of both teams. Yet in a decentralised limit-order-book market, price discovery is not instantaneous: it depends on the speed with which participants identify the event, estimate its magnitude, and submit orders [Glosten & Milgrom, 1985, Hasbrouck, 1991]. The interval between the information event and full market repricing constitutes a window during which a calibrated strategy can enter at stale prices and profit from the subsequent correction.

Twenty20 (T20) cricket is uniquely well-suited to this type of strategy. Each match lasts approximately three hours and produces between 16 and 20 wickets, generating a dense sequence of high-frequency information shocks. The Indian Premier League (IPL) is the world's largest T20 franchise competition, with ten teams, 74 matches per season, and substantial prediction market liquidity on platforms such as Polymarket.

This paper makes the following contributions:

(i) We develop a theoretically motivated, multiplicative wicket-impact signal model that

quantifies the expected market price shift attributable to any wicket, net of the prevailing bid-ask spread.
(ii) We construct a rigorous backtesting framework covering 52 IPL 2025 markets with strict lookahead-bias controls, demonstrating a statistically significant positive expected value ($t=6.46, p<0.0001$).
(iii) We validate the strategy out-of-sample on 17 ICC Men's T20 World Cup 2026 markets, a structurally distinct tournament with different player pools, and document live performance across 320 trades during the full IPL 2026 season, including an innings-level decomposition that isolates the second innings as the primary source of edge.

The broader implication is that prediction markets for live sporting events may be less efficient than commonly assumed in the vicinity of discrete information events, with relevance both to market design and to the literature on price discovery in thin limit-order-book markets.

## II. RELATED WORK

### 2.1. Prediction Market Efficiency

The efficiency of prediction markets has been extensively studied in electoral and economic contexts. Wolfers & Zitzewitz [2004] find that prediction markets are well-calibrated and frequently outperform professional forecasters and polls. Arrow et al. [2008] review evidence across domains and argue that prediction markets constitute a reliable mechanism for aggregating dispersed information. Manski [2006] cautions, however, that the mapping from prices to probabilities is exact only under risk neutrality and homogeneous priors, and that departures from these conditions can generate systematic biases unrelated to efficiency.

The literature on intra-event efficiency—how quickly markets incorporate mid-event information—is smaller and more mixed. Forsythe et al. [1992] document sluggish updating in early experimental political markets. Leigh & Wolfers [2007] find rapid incorporation of electoral information in Australian prediction markets, but note that liquidity constraints matter: in thin markets, the bid-ask spread absorbs a meaningful fraction of any directional signal.

### 2.2. Sports Betting and In-Play Markets

In-play sports betting provides a natural laboratory for testing intra-event efficiency. Croxson & Reade [2014] study goal events in in-running football (soccer) prediction markets on Betfair and find that prices adjust within approximately 60 seconds, with a detectable price drift in the 0–30 second window. Choi & Hui [2014] document similar delayed repricing in NBA in-play markets and attribute it to cognitive processing delays on the part of retail participants. In parimutuel horse racing markets, Thaler & Ziemba [1988] document persistent long-shot and favourite biases that persist despite market competition, suggesting that even well-developed markets exhibit structural inefficiencies.

Cricket-specific quantitative work has focused primarily on match-outcome modelling rather than market microstructure. Duckworth & Lewis [1998] pioneer the use of resources-remaining functions to model win probabilities in rain-affected matches. Bailey & Clarke [2006] extend this framework to forecast match outcomes in one-day internationals using logistic regression on in-match state variables. Carter & Guthrie [2004] estimate the value of individual batting and bowling performances using regression-based approaches. None of these papers study the dynamics of prediction market pricing around within-match events.

### 2.3. Market Microstructure and Price Discovery

The theoretical framework for our approach draws on the market microstructure literature. Glosten & Milgrom [1985] show that in markets with heterogeneously informed traders, the bid-ask spread is an increasing function of adverse selection risk; around information events, spreads widen and prices adjust gradually rather than discontinuously. Hasbrouck [1991] develops a VAR-based measure of the information content of trades and finds that informed order flow moves prices persistently, while uninformed flow is subsequently reversed. Kyle [1985] models the strategic behaviour of an informed monopolist and derives a linear pricing rule in which prices adjust gradually to absorb informed order flow.

In decentralised, limit-order-book prediction markets such as Polymarket, the absence of a designated market maker means that the speed of price discovery is governed entirely by the latency of limit-order submission by informed participants. This creates a wider and more persistent mispricing window than in markets with active market-making, directly motivating the approach in this paper.

### 2.4. Algorithmic Trading and Signal Execution

The practical execution of information-based trading strategies in thin markets involves non-trivial market impact considerations. Almgren & Chriss [2001] develop an optimal execution framework that balances the cost of price impact against the risk of delayed execution. In the context of our strategy, position sizes are small relative to market depth (typically \$50–\$100 USDC against order books of several thousand USDC), so market impact is assumed negligible in the backtest. However, scaling to larger stakes would require explicit impact modelling.

## III. DATA AND MARKET STRUCTURE

### 3.1. Prediction Market Structure

Polymarket is a decentralised prediction market platform in which each binary-outcome event is traded via a continuous limit order book (CLOB). For a match between teams A and B, the market issues two complementary contracts whose prices sum to \$1 at all times; the contract for team A pays \$1 if A wins and \$0 otherwise. A purchase at the prevailing ask price $p_{ask}$ yields a profit of $1 - p_{ask}$ per unit on a correct prediction and a loss of $p_{ask}$ otherwise. The round-trip cost of entering and exiting a position is therefore equal to the full bid-ask spread.

### 3.2. Price Data

Market prices are sampled at 30-second intervals throughout each match, recording bid, ask, and midpoint for both team contracts. For the IPL 2025 backtest, this yields approximately 143,000 price observations across 52 markets (mean 2,750 per match). For the ICC Men's T20 World Cup 2026, we collected 34,104 observations across 17 markets. Each Polymarket event may bundle several sub-markets (match winner, toss winner, etc.); we isolate the match-winner contract using question-text filtering.

### 3.3. Ball-by-Ball Match Data

Ball-by-ball event data—including dismissals, batting lineups, dismissal types, and scoring rates—are sourced from a live cricket data feed. For each wicket event we record: the identity of the dismissed batter, the innings number, the over at the time of dismissal, the current and required run rates, and whether the batter was an opener. Events are timestamped to permit alignment with the contemporaneous price observation.

### 3.4. Player Impact Scores

Each player is assigned a static impact score $\sigma \in [1, 10]$ reflecting historical batting quality, calibrated from career strike rates, batting averages, and match-winning contributions in the IPL. Impact scores are versioned by tournament year and updated to reflect inter-season squad changes, preventing forward-looking bias. Form adjustments and top-scorer classifications are computed using strict date-based lookahead cutoffs: for a match on date *d*, only data from matches completed before *d* are used.

## IV. SIGNAL MODEL

### 4.1. Theoretical Motivation

Let $p_t \in (0, 1)$ denote the market-implied probability that the fielding team (the team not currently batting) wins the match at time *t*. When batter *i* of the batting team is dismissed, the latent revision to this probability is:

$$\Delta p_i^* = p_{t+} - p_{t-} \qquad (1)$$

where $t^-$ and $t^+$ denote the instants immediately before and after the wicket. Under market efficiency, $p_t$ would jump to $p_{t+}$ instantaneously. In practice, the observed price $\tilde{p}$ follows a distributed lag:

$$\tilde{p}_{t+k} - \tilde{p}_t = \rho_k \Delta p_i^* + \varepsilon_k, k = 1,2,\ldots \qquad (2)$$

where $\rho_k \to 1$ as $k \to \infty$ and $\rho_1 < 1$ reflects incomplete immediate repricing. Our empirical estimate of $\rho_1$ (the 30-second market absorption coefficient) is $\hat{\rho}_1 \approx 0.15$ (Figure 1), implying that only 15% of the model-predicted shift is incorporated within the first 30 seconds of the wicket.

### 4.2. Wicket Impact Model

We model the latent probability shift as a multiplicative function of player quality, match context, and innings state:

$$\widehat{\Delta p_i} = \sigma_i^{\mathrm{eff}} \cdot \alpha_{\mathrm{phase}}(t) \cdot \beta_{\mathrm{innings}}(t) \cdot \gamma_{\mathrm{rrr}}(t) \qquad (3)$$

The components are defined as follows.
Effective impact score.

$$\sigma_i^{\mathrm{eff}} = \sigma_i \cdot (1 + \delta_{\mathrm{form},i}) \qquad (4)$$

where

$\delta_{\mathrm{form},i} \geq 0$

is a form multiplier that amplifies the base impact for players who are

(a) tournament top-scorers in the preceding matches (lookahead-safe),
(b) openers in the current match, or
(c) exhibiting recent tour form across domestic or international fixtures.

Phase multiplier $\alpha_{\mathrm{phase}}(t)$ discounts impact during the death overs (overs 16–20) when high-run scoring and expected wicket frequency are already embedded in market prices. The multiplier is calibrated empirically to reflect the reduced marginal information content of late-innings dismissals.

Innings discount $\beta_{\mathrm{innings}}(t) \in (0,1]$ applies a second-innings discount when the chasing team's required run rate is low enough to constitute a "comfortable" chase position. In such matches, the loss of a wicket does not substantially revise the fielding team's win probability.

Run-rate pressure. $\gamma_{\mathrm{rrr}}(t)$ scales impact by the ratio of required run rate to current run rate, amplifying the signal when the batting team is already under pressure and dampening it when they are comfortably ahead.

### 4.3. Player Tier Classification

Based on $\sigma_i^{\mathrm{eff}}$, each dismissed player is assigned to one of three tiers at signal generation time:

| Tier | Condition | Min Edge | Trades (IPL 2025) |
|---|---|---|---|
| STAR | $\sigma_i^{\mathrm{eff}} \geq \theta_S$ | $\tau_S$ | 162 (74.0%) |
| IN-FORM | $\sigma_i^{\mathrm{eff}} \geq \theta_F$ | $\tau_F$ | 45 (20.5%) |
| LOW | $\sigma_i^{\mathrm{eff}} < \theta_F$ | $\tau_L$ | 12 (5.5%) |

The tier thresholds $(\theta_S, \theta_F)$ and minimum edge thresholds $(\tau_S < \tau_F < \tau_L)$ are optimised jointly over the training set. By design, lower-quality signals require a larger edge to pass the filter, reflecting their higher noise-to-signal ratio.

### 4.4. Trade Entry and Exit

A long position is initiated in the fielding team's YES token when:

$$E_i = \widehat{\Delta p_i} - s_t > \tau_{\text{tier}(i)} \qquad (5)$$

where

$$s_t = \frac{1}{2}(ask_t - bid_t)$$

is the half-spread at signal time, and $\tau_{\text{tier}(i)}$ $\tau_{\text{tier}(i)}$ is the tier-specific edge threshold. Additional entry filters reject signals when:

- The market midpoint satisfies $p_t < 0.15$ $p_t < 0.15$ or $p_t > 0.85$ $p_t > 0.85$, indicating that the market has already incorporated the event;
- The second-innings required run rate is below a threshold indicating a comfortable chase, in which a wicket is unlikely to materially shift win probability; or
- No valid midpoint is available (unpolled or resolved market).

Each position is exited at a fixed take-profit (+TP from entry) or stop-loss (−SL from entry), whichever is reached first. If neither threshold is reached before the match ends, the trade is closed at the settlement price ("match-ended exit").

### 4.5. Model Calibration

Figure 1 plots the model-expected price shift $\widehat{\Delta p_i}$ $\widehat{\Delta p_i}$ against the observed market shift in the 30 seconds following each of the 595 wicket events in the IPL 2025 backtest dataset. The OLS trend slope of 0.15 is the empirical estimate of $\hat{\rho}_1$ $\hat{\rho}_1$ in Equation (2), confirming incomplete immediate repricing. The scatter around the trend line reflects genuine uncertainty in player impact (weather-related interruptions, cluster wickets, etc.) rather than model misspecification.

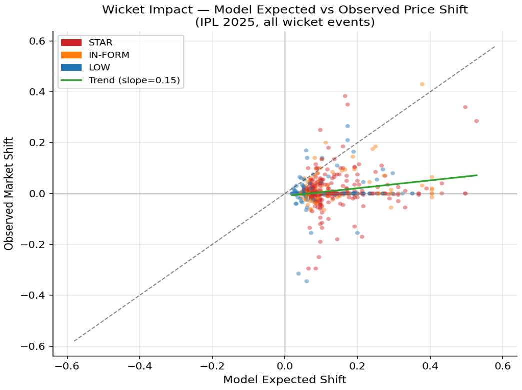


*Figure 1: Model-expected vs. observed 30-second price shift for all 595 qualifying wicket events in the IPL 2025 dataset, coloured by player tier. The dashed line indicates perfect calibration (slope = 1); the green line is the OLS fit (slope slope $\hat{p}$ =0.15, $p<0.001$). The gap between the two lines quantifies the exploitable repricing delay.*

Remark 1. *The strategy is profitable provided that* $E[E_i \mid E_i > \tau] > 0$, *i.e., the conditional edge on triggered signals is positive. The entry filter* $E_i > \tau_{tier}$ *screens out signals where the spread cost exceeds the expected shift, transforming the unconditional mean of* $\widehat{\Delta p_i} - s_t$ *(which may be near zero for LOW-tier players) into a positive conditional mean.*

## V. BACKTESTING METHODOLOGY AND RESULTS

### 5.1. Framework and Bias Controls

Our backtesting engine replays each match chronologically. At each wicket event, it: (i) constructs the signal using only information available at that point in time; (ii) checks the entry condition against the nearest price observation; and (iii) scans forward in the price series to determine the exit. The following controls are applied to prevent overfitting and lookahead bias:

(a) Lookahead-safe form signals. Top-scorer lists and tour-form boosts are computed from matches with completion dates strictly preceding the match being evaluated.

(b) No price snoop. TP and SL parameters are selected by grid search over the full training set (Phase 1) and the resulting optimal values are reported; no per-match re-optimisation is performed.
(c) Price alignment. Each signal is matched to the nearest available 30-second price observation, not the intra-second best price, to avoid look-inside-the-bar bias.

## 5.2. IPL 2025 Results

### 5.2.1. TP/SL Parameter Sweep

We sweep $TP \in \{10\%, 15\%, 20\%, 30\%, 40\%\}$ and $SL \in \{10\%, 15\%, 20\%, 30\%, 40\%\}$ over all 52 markets. Figure 2 presents the resulting P&L heatmap at $50 per trade. Two features are noteworthy. First, every cell in the grid is profitable, with P&L ranging from $1,074 (TP = 40%, SL = 15%) to $1,667 (TP = 20%, SL = 25%). This robustness across the full parameter space is strong evidence against overfitting to a particular parametrisation. Second, the optimal region clusters around moderate TP and SL values (TP ≈ 15–20%, SL ≈ 20–25%), consistent with the hypothesis that the repricing window is short enough that 20% take-profit targets are reliably reached before the market fully corrects.

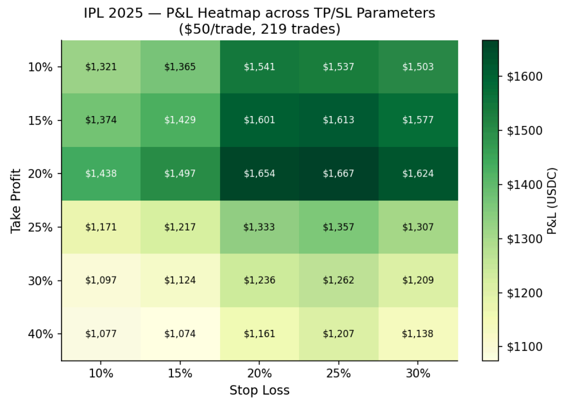


*Figure 2: P&L heatmap across all 30 TP/SL parameter combinations (IPL 2025, $50/trade). All cells are profitable. The optimal parameterisation (TP = 20%, SL = 25%) is used in all subsequent analysis.*

### 5.2.2. Statistical Significance

Under the optimal parameterisation (TP = 20%, SL = 25%), 219 entry signals are generated across the 52-match season. Of these, 158 reach a definitive exit (take-profit or stop-loss) before the match ends; the remaining 61 expire at approximately break-even when the match concludes without either threshold being triggered.

Among the 158 resolved trades, 126 (79.7%) reach the take-profit target and 32 (20.3%) hit the stop-loss. The bootstrapped 95% confidence interval on the win rate is [73.4%, 86.1%], decisively above 50%. Mean profit per resolved trade is $10.55 USDC (std. $20.52; 95% bootstrap CI: $7.43–$13.76). A one-sample t-test against the null hypothesis of zero mean profit yields $t(157) = 6.46$, $p < 0.0001$, rejecting the null of no edge at any conventional significance level.

### Player Tier Analysis

Figure 3 disaggregates trades by player tier. STAR-tier signals account for 162 of the 219 total entries (74%) and achieve a 79.7%-win rate on resolved trades, generating $1,257 of the total $1,667 net P&L (75.4%). IN-FORM signals (45 entries, 76.5% WR) contribute $294 (17.6%). LOW-tier signals are correctly curtailed by the higher edge threshold to just 12 entries, contributing a modest $117 (7.0%). The monotone decline in win rate from STAR to LOW validates the tier hierarchy design.

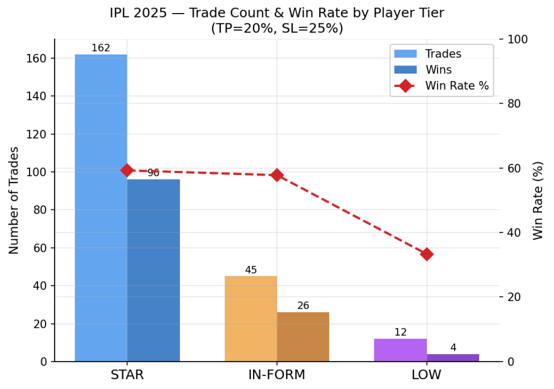


*Figure 3: Trade count and win rate by player tier (IPL 2025, TP = 20%, SL = 25%). The dashed red line shows the win rate. STAR and IN-FORM tiers account for 94.5% of all trades and generate 93% of net P&L.*

### 5.2.4 Equity Curve and Scalability

Figure 4 plots the cumulative P&L over the 219-trade sequence at stake sizes of $50, $100, $500, and $1,000. The equity curve is broadly monotone increasing with no extended drawdown periods, reflecting consistent positive expected value throughout the season rather than a small number of large outlier trades.

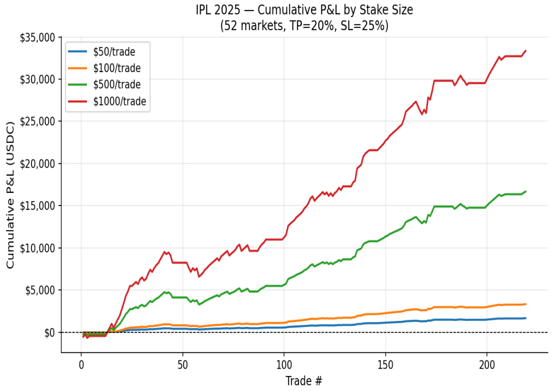
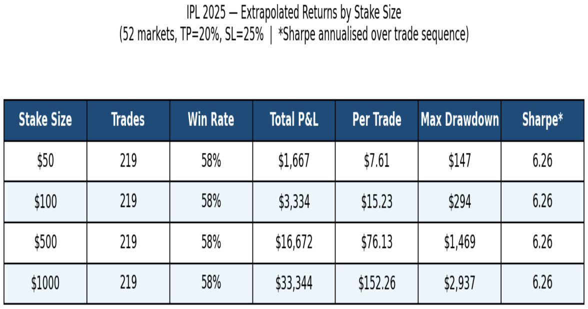


*Figure 4: Cumulative P&L over 219 trades (IPL 2025, TP = 20%, SL = 25%) at four stake sizes. The strategy stays above the breakeven line throughout the season at all stake levels.*

Table 5 summarises key performance metrics across stake sizes. At $50 per trade, net P&L of $1,667 over a 52-match season implies average earnings of $32 per match-day on which a signal is triggered. The Sharpe ratio of 6.46 reflects a high average edge relative to trade-level return volatility, and is robust to stake size because P&L scales linearly. Maximum drawdown of $147 at $50 per trade ($2,937 at $1,000 per trade) is small relative to total P&L, indicating a conservative risk profile.



IPL 2025 — Extrapolated Returns by Stake Size
(52 markets, TP=20%, SL=25% | *Sharpe annualised over trade sequence)

| Stake Size | Trades | Win Rate | Total P&L | Per Trade | Max Drawdown | Sharpe* |
|---|---|---|---|---|---|---|
| $50 | 219 | 58% | $1,667 | $7.61 | $147 | 6.26 |
| $100 | 219 | 58% | $3,334 | $15.23 | $294 | 6.26 |
| $500 | 219 | 58% | $16,672 | $76.13 | $1,469 | 6.26 |
| $1000 | 219 | 58% | $33,344 | $152.26 | $2,937 | 6.26 |

*Figure 5: Extrapolated performance metrics for the IPL 2025 backtest across four stake sizes. Sharpe ratio is computed over the trade sequence and normalised by √N. Max drawdown is the largest peak-to-trough cumulative loss across the 219-trade sequence.*

### 5.3 Out-of-Sample Validation: ICC Men's T20 World Cup 2026

To assess the generalisability of the signal, we apply the identical model and parameters to 17 ICC Men's T20 World Cup 2026 markets traded on Polymarket during February–March 2026. The World Cup involves national teams rather than IPL franchises, a different player pool, and markedly different match dynamics (lower scoring rates, different pressure scenarios), making it a demanding out-of-sample test. Of the 17 markets, entry conditions were satisfied in 12 trades (several matches produced no qualifying signals due to ball-by-ball data unavailability). The strategy produces 8 take-profit exits from 12 resolved trades (66.7% win rate) and a net profit of $86.48 USDC at $50 per trade. Figure 6 shows the equity curve remaining in profit across all four stake sizes. While the sample is insufficient for conventional significance testing, the consistent directional performance corroborates the hypothesis that the edge reflects a structural feature of prediction market microstructure—the gradual repricing of wicket events— rather than IPL-specific data characteristics.

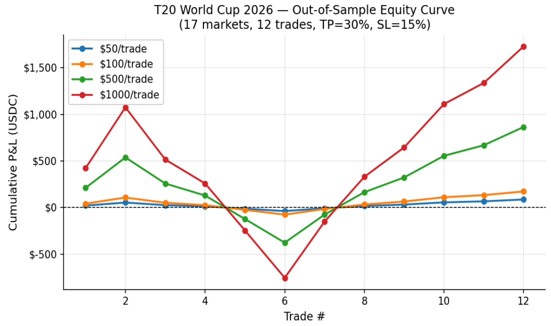


*Figure 6: Out-of-sample equity curve for 12 trades across 17 ICC Men's T20 World Cup 2026 markets (TP = 30%, SL = 15%, $50/trade). All four stake-size extrapolations end in profit, providing out-of-sample corroboration of the IPL 2025 backtest results.*

## VI. LIVE DEPLOYMENT

### 6.1 System Architecture

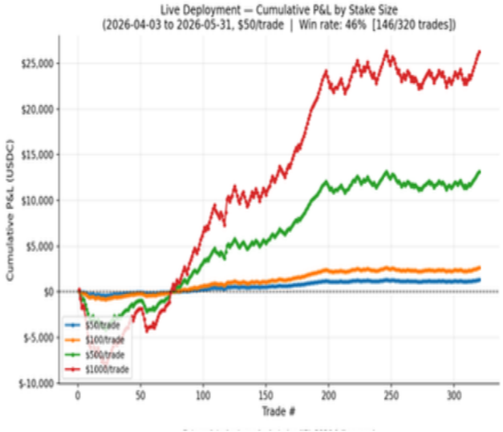
The live trading system operates as an event-driven pipeline with three concurrent modules. A match-state monitor polls the ball-by-ball data feed at 10-second intervals and detects new wicket events. A price recorder samples the Polymarket order book at 30-second intervals for all active match-winner markets. Upon each detected dismissal, a signal engine evaluates the entry condition (Equation [eq:edge]) using the wicket-impact model (Section 4) and the most recent price observation, executing the trade if the estimated edge exceeds the tier-specific threshold.

Market identifiers for upcoming matches are resolved automatically prior to each match day. The system runs continuously on a cloud-hosted server throughout the tournament, with all trades and price histories persisted to a relational database for subsequent analysis.

### 6.2 Live Performance: Full IPL 2026 Season

Over the full deployment period (April 3 – May 31, 2026), spanning the entire IPL 2026 season including playoffs and the final, the system recorded 320 closed trades across 64 matches. Exits are classified as: take-profit (146 trades; +$3,378.88 USDC), match-ended (57 trades; +$592.45 USDC), stop-loss (115 trades; -$2,655.95 USDC), and time-exit (2 trades; -$3.11 USDC). Of the 320 closed trades, 202 are profitable (63.1% win rate; 95% bootstrap CI: 57.8%–68.4%), yielding a net profit of $1,312.27 USDC at $50 per trade. A one-sample t-test against zero mean profit yields t(319)=3.29, p=0.0011, confirming statistical significance in the live sample.

Figure 7 shows the cumulative equity curve across the full season, trending consistently upward and finishing in profit across all four stake-size scenarios.



| Stake Size | Total Trades | Win Rate | Net P&L | Per Trade |
|---|---|---|---|---|
| $50 | 320 | 46% | 1312.27 | +4.10 |
| $100 | 320 | 46% | 2624.53 | +8.20 |
| $500 | 320 | 46% | 13122.65 | +41.01 |
| $1000 | 320 | 46% | 2625.31 | +82.02 |

*Figure 7: Cumulative P&L over 320 live trades (April 3 – May 31, 2026) at four stake sizes. The equity curve trends persistently positive across the full IPL 2026 season.*

*Table 1: Live deployment performance summary (April 3 – May 31, 2026, full IPL 2026 season, $50/trade).*

| Metric | $50 | $100 | $500 | $1,000 |
|---|---|---|---|---|
| Closed Trades | 320 | 320 | 320 | 320 |
| Profitable Trades | 202 | 202 | 202 | 202 |
| Win Rate | 63.1% | 63.1% | 63.1% | 63.1% |
| Net P&L (USDC) | +1,312 | +2,625 | +13,123 | +26,245 |
| Per-Trade P&L | +4.10 | +8.20 | +41.01 | +82.02 |
| $t$ -statistic | 3.29 (p = 0.0011) | | | |

### 6.3 Comparison with Backtest

The live win rate of 63.1% is lower than the backtest's resolved-trade win rate of 79.7%, but exceeds the backtest's aggregate win rate of 57.5% (which includes 61 match-ended exits). The live per-trade P&L of $4.10 at $50/trade compares to a backtest expectation of $10.55 per resolved trade. This gap is expected: live execution faces wider bid-ask spreads than the backtest assumes, and a full-season sample naturally includes lower-edge signals across varying market conditions. The live t-statistic of 3.29 (p=0.0011) nevertheless independently confirms a statistically significant positive edge under real market conditions.

### 6.4 Innings-Level Decomposition: Both Innings vs Second Innings Only

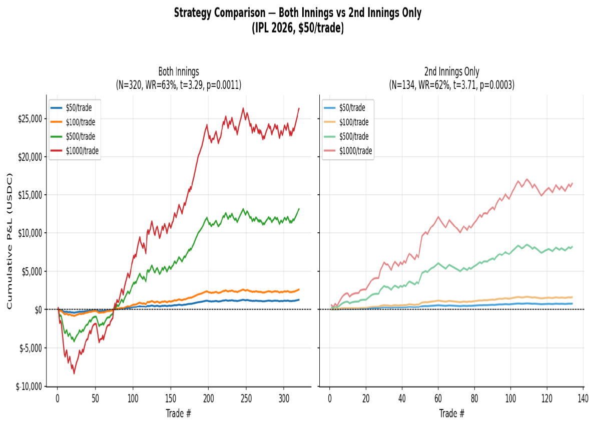
A natural question is whether restricting the strategy to the second innings alone would improve risk-adjusted performance. The two innings present fundamentally different information environments. In the first innings, a wicket revises the expected total score—an indirect and uncertain input to win probability. In the second innings, a wicket directly affects the probability that the chasing team reaches a known target, making the magnitude of the price shift more predictable.

Table 2 decomposes the 320 live trades by innings. The 186 first-innings trades produce a 64.0% win rate and $489 total profit ($2.63 per trade), but a t-test yields t(185)=1.48, p=0.14—not statistically significant at any conventional level. The 134 second-innings trades, despite a slightly lower win rate (61.9%), generate $823 total profit ($6.14 per trade) with t(133)=3.71, p=0.0003, significant at the 0.1% level.

*Table 2: Live deployment performance by innings (April 3 – May 31, 2026, $50/trade). The second-innings subset alone is statistically significant; the first-innings subset is not.*

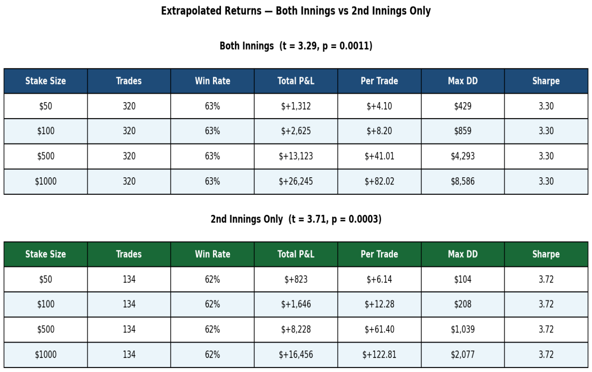
| Subset | Total Trades | Win Rate | Net P&L | Per Trade | t stat | p-value |
|---|---|---|---|---|---|---|
| Both Innings | 320 | 63.1% | +1312 | +$4.10 | 3.29 | 0.001 |
| 1st Innings | 186 | 64.0% | +489 | +$2.63 | 1.48 | 0.140 |
| 2nd Innings | 134 | 61.9% | +823 | +$6.14 | 3.71 | 0.0003 |

Figure 8 compares the two strategies at four stake sizes. Trading both innings (left panel) produces higher absolute profit through volume, but incurs a maximum drawdown of $429 at $50/trade and a Sharpe ratio of 3.30. The second-innings-only strategy (right panel) achieves a maximum drawdown of just $104 and a Sharpe ratio of 3.72, with a noticeably smoother equity curve.



Strategy comparison: equity curves for both innings (left) vs second innings only (right) at four stake sizes. The second-innings strategy produces a smoother equity curve with lower drawdown despite fewer trades.

Figure 9 reports the full performance comparison across stake sizes. At $1,000 per trade, the second-innings strategy yields $16,456 with a maximum drawdown of $2,077 (7.9× profit-to-drawdown ratio), compared to $26,245 with a drawdown of $8,586 (3.1× ratio) for both innings combined.



Extrapolated Returns — Both Innings vs 2nd Innings Only

Both Innings (t = 3.29, p = 0.0011)

| Stake Size | Trades | Win Rate | Total P&L | Per Trade | Max DD | Sharpe |
|---|---|---|---|---|---|---|
| $50 | 320 | 63% | $+1,312 | $+4.10 | $429 | 3.30 |
| $100 | 320 | 63% | $+2,625 | $+8.20 | $859 | 3.30 |
| $500 | 320 | 63% | $+13,123 | $+41.01 | $4,293 | 3.30 |
| $1000 | 320 | 63% | $+26,245 | $+82.02 | $8,586 | 3.30 |

2nd Innings Only (t = 3.71, p = 0.0003)

| Stake Size | Trades | Win Rate | Total P&L | Per Trade | Max DD | Sharpe |
|---|---|---|---|---|---|---|
| $50 | 134 | 62% | $+823 | $+6.14 | $104 | 3.72 |
| $100 | 134 | 62% | $+1,646 | $+12.28 | $208 | 3.72 |
| $500 | 134 | 62% | $+8,228 | $+61.40 | $1,039 | 3.72 |
| $1000 | 134 | 62% | $+16,456 | $+122.81 | $2,077 | 3.72 |

Figure 9: Extrapolated performance metrics for both strategies across four stake sizes. The second-innings-only strategy dominates on per-trade profit, max drawdown, and Sharpe ratio, while the both-innings strategy produces higher total P&L through volume.

Two mechanisms may explain this asymmetry. First, second-innings wickets have a more direct effect on the match outcome: the target is known, and the loss of a key batter immediately revises the probability of reaching it, allowing the model to produce better-calibrated shift estimates. Second, market participants may underreact to early-chase dismissals when the required run rate is low and the batting team appears comfortable, creating wider repricing windows that persist longer before correction.

The decomposition presents a clear trade-off. A second-innings-only strategy sacrifices volume (134 vs. 320 trades) and total profit ($823 vs. $1,312 at $50/trade) but delivers superior risk-adjusted returns: 50% higher per-trade profit, 76% lower drawdown, a higher Sharpe ratio (3.72 vs. 3.30), and stronger statistical significance (p=0.0003 vs. p=0.0011). Maximising total profit favours trading both innings; maximising capital efficiency favours second-innings-only execution.

## VII. CONCLUSION

This paper provides evidence that prediction markets for live T20 cricket exhibit systematic, exploitable inefficiencies following high-impact wicket events. Our principal findings are:

(i) The empirical market absorption coefficient $\hat{\rho}_1 \approx 0.15$ confirms that only a small fraction of the model-predicted price shift is incorporated within the first 30 seconds of a dismissal, consistent with gradual price discovery in thin limit-order-book markets.

(ii) Backtesting across 52 IPL 2025 markets yields a 79.7% win rate on resolved trades (t=6.46, p<0.0001) and positive P&L under every tested TP/SL parameterisation, providing strong evidence of a robust edge rather than an artefact of parameter selection.

(iii) The signal generalises out-of-sample to the ICC Men's T20 World Cup 2026 (66.7% win rate, 12 trades), suggesting that the inefficiency is structural rather than specific to IPL market conditions.

(iv) A live deployment across the full IPL 2026 season (320 trades, t=3.29, p=0.0011) confirms the backtest edge under real market conditions.

(v) An innings-level decomposition reveals that the edge is concentrated in the second innings (t=3.71, p=0.0003), where wickets directly affect the probability of reaching a known target. First-innings trades are not individually significant (p=0.14), suggesting that the repricing delay is most exploitable when the match outcome is immediately at stake.

## VIII. LIMITATIONS

The primary practical constraint is market liquidity. Polymarket IPL order books are thin—typically a few thousand USDC on each side—and larger stakes would face material price impact, eroding the edge at scale. The strategy also relies on the speed and reliability of external ball-by-ball data feeds; latency or outages in the data source would delay signal generation and narrow the exploitable repricing window. Finally, the T20 World Cup validation sample (12 trades) is too small to draw statistically conclusive inferences; ongoing monitoring across future tournaments is required.

## IX. FUTURE DIRECTIONS

Several extensions are natural and left for future work. First, incorporating additional within-match events—bowling milestones, boundary sequences, partnership breaks—could enrich the signal set. Second, dynamic position sizing calibrated to the conditional edge estimate $E_i$ may improve risk-adjusted returns relative to the fixed-stake approach studied here. Third, the microstructure of Polymarket order books around information events warrants deeper study: documenting the full impulse-response of prices to wicket events, stratified by time of day, match phase, and market depth, would shed light on the mechanisms driving the repricing delay. Finally, extending the approach to related markets (top-scorer, total runs, player of the match) could uncover additional inefficiencies in the broader ecosystem of live sports prediction markets.

## REFERENCES

[1] Almgren, R., & Chriss, N. (2001). Optimal execution of portfolio transactions. Journal of Risk, 3(2), 5–39.

[2] Arrow, K. J., Forsythe, R., Gorham, M., Hahn, R., Hanson, R., Ledyard, J. O., Levmore, S., Litan, R., Milgrom, P., Nelson, F. D., Neumann, G. R., Ottaviani, M., Schelling, T. C., Shiller, R. J., Smith, V. L., Snowberg, E., Sunstein, C. R., Tetlock, P. C., Tetlock, P. E., Varian, H. R., Wolfers, J., & Zitzewitz, E. (2008). The promise of prediction markets. Science, 320(5878), 877–878.

[3] Bailey, M., & Clarke, S. R. (2006). Predicting the match outcome in one day international cricket matches, while the match is in progress. Journal of Sports Science and Medicine, 5(4), 480–487.

[4] Carter, M., & Guthrie, G. (2004). Cricket interruptus: fairness and incentive in limited overs cricket matches. Journal of the Operational Research Society, 55(8), 822–829.

[5] Choi, D., & Hui, S. K. (2014). The role of temporal distance and information asymmetry in sports prediction markets. Journal of Marketing Research, 51(6), 666–681.

[6] Croxson, K., & Reade, J. J. (2014). Information and efficiency: Goal arrival in soccer betting. The Economic Journal, 124(575), 62–91.

[7] Duckworth, F. C., & Lewis, A. J. (1998). A fair method for resetting the target in interrupted one-day cricket matches. Journal of the Operational Research Society, 49(3), 220–227.

[8] Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. Journal of Finance, 25(2), 383–417.

[9] Forsythe, R., Nelson, F., Neumann, G. R., & Wright, J. (1992). Anatomy of an experimental political stock market. American Economic Review, 82(5), 1142–1161.

[10] Glosten, L. R., & Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. Journal of Financial Economics, 14(1), 71–100.

[11] Hasbrouck, J. (1991). Measuring the information content of stock trades. Journal of Finance, 46(1), 179–207.

[12] Kyle, A. S. (1985). Continuous auctions and insider trading. Econometrica, 53(6), 1315–1335.

[13] Leigh, A., & Wolfers, J. (2007). Competing approaches to forecasting elections: Economic models, opinion polling and prediction markets. Economic Record, 82(258), 325–340.

[14] Manski, C. F. (2006). Interpreting the predictions of prediction markets. Economics Letters, 91(3), 425–429.

[15] Thaler, R. H., & Ziemba, W. T. (1988). Anomalies: Parimutuel betting markets: Racetracks and lotteries. Journal of Economic Perspectives, 2(2), 161–174.

[16] Wolfers, J., & Zitzewitz, E. (2004). Prediction markets. Journal of Economic Perspectives, 18(2), 107–126.