Are Final Market Prices Sufficient for Information Aggregation? Evidence from Last-Minute Dynamics in Parimutuel Betting*

Hiroaki Hanyu† Shunsuke Ishii‡ Suguru Otani§ Kazuhiro Teramoto¶

September 18, 2025

Abstract

Most betting market models employ static frameworks that condition decisions on final odds. Using a unique dataset of interim odds from Japanese horse racing, this study examines the validity of such static analyses by asking whether there is a systematic relationship between expected returns and the trajectory of odds. We find that returns are negatively related to last-minute changes in odds, and that these late movements attenuate the favorite-longshot bias by weakening the correlation between final odds and returns. These patterns suggest that informed bettors place wagers at the final stage based on private signals, leaving surprises in final odds.

Keywords: Information aggregation, Parimutuel mechanism, Betting markets, Favorite-longshot bias

JEL code: G14, D47, D83, L83

1 Introduction

Parimutuel betting markets offer a well-defined empirical setting for analyzing equilibrium behavior under uncertainty. One of the most enduring puzzles in these markets is the *favorite-longshot bias*

*The authors are grateful to Pierre-André Chiappori, Nobuyuki Hanaki, Yuichiro Kamada, Hayato Kanayama, Michihiro Kandori, Shunya Noda, Yasuhiro Shirata, and Ryoji Yanashima for their insightful comments and suggestions. We also thank the participants of the SWET2025 Market Design session for helpful feedback and discussions. This research uses data from the JRA-VAN database (a licensed service providing structured JRA data), with official permission granted by JRA-VAN for academic use. Financial support from JST ERATO Grant Number JPMJER2301 is gratefully acknowledged.

†bd251004@g.hit-u.ac.jp, Hitotsubashi University

‡s.ishii@lse.ac.uk, London School of Economics and Political Science

§suguru.otani@e.u-tokyo.ac.jp, Market Design Center, Graduate School of Economics, University of Tokyo

¶k.teramoto@r.hit-u.ac.jp, Graduate School of Economics, Hitotsubashi University.

(FLB), first documented by Griffith (1949). The FLB refers to the systematic tendency for wagers on longshots—horses with high odds and low perceived winning probabilities—to generate lower average returns than wagers on favorites—horses with low odds and high perceived chances of success. This pattern stands in sharp contrast to conventional notions of market efficiency (Thaler and Ziemba 1988). A large body of research has attempted to account for this anomaly through equilibrium models that link final odds to win probabilities (e.g., Weitzman 1965, Ali 1977, Hurley and McDonough 1995, Jullien and Salanié 2000, Bradley 2003, Snowberg and Wolfers 2010).[1] More recent studies, including Gandhi and Serrano-Padial (2015) and Chiappori, Salanié, Salanié and Gandhi (2019), develop models with heterogeneous bettors and use structural estimation to recover risk preferences, probability perceptions, and subjective beliefs from the relationship between final odds and realized outcomes.

This study is motivated by the central question whether the *static* relationship between final odds and realized outcomes is sufficient for identifying bettors' risk preferences and beliefs. While existing studies have substantially advanced our understanding of pricing in betting markets, they largely rely on static frameworks in which bettors are assumed to condition their choices on final odds. However, this assumption is questionable as parimutuel betting markets, unlike financial markets where limit orders are feasible, do not allow bettors to act contingent on final odds. In such markets, strategic delays by informed bettors in placing their wagers can prevent efficient information aggregation and thereby generate the FLB, independently of bettors' risk preferences or probability perceptions (Ottaviani and Sørensen 2009).

To address our research question, we empirically test the validity of inferences that rest on the assumption that final odds serve as sufficient statistics for bettors' decisions. More specifically, we analyze the entire evolution of odds throughout the betting period to assess whether outcomes depend not only on final odds but also on the trajectory by which those odds are reached. If final odds are truly sufficient statistics, then horses with identical final odds should, on average, yield identical expected returns regardless of the path of odds movements. Any systematic deviation from this benchmark would indicate that odds dynamics contain predictive information beyond what final odds alone can capture.

For our analysis, we employ a comprehensive dataset from the Japan Racing Association (JRA) that covers all centrally administered races from 2004 to 2023. A distinctive feature of this dataset is the availability of high-frequency interim odds, recorded at five-minute intervals from the opening of the betting window until one minute before post time. These interim observations allow us to trace the real-time evolution of market expectations and to examine whether, and in what ways, the

[1] A comprehensive review of empirical work estimating risk preferences using field data is provided by Barseghyan, Molinari, O'Donoghue and Teitelbaum (2018), who highlight betting markets as a key setting for identifying such preferences from aggregate behavior. See also the surveys by Thaler and Ziemba (1988), Hausch and Ziemba (1995), Ottaviani and Sørensen (2008), and Jullien and Salanié (2008). For experimental and theoretical perspectives, see Plott, Wit and Yang (2003), Koessler, Noussair and Ziegelmeyer (2012), Kajii and Watanabe (2017), and Gillen, Plott and Shum (2017).

trajectory of odds—particularly in the final stages of betting—affects subsequent realized returns.

In parimutuel betting markets, final odds and interim odds are calculated in the same way—based on the cumulative distribution of wagers—but they serve markedly different functions. Final odds are determined at the close of betting and directly govern payout calculations, thereby playing a dual role as both a summary of aggregate market expectations and as payoff-determining prices. Interim odds, by contrast, are snapshots of cumulative bets at intermediate points. Although they may capture evolving expectations, they do not enter the payout rule and thus lack a direct pricing function. Because of this, conventional models of betting markets have typically abstracted from interim odds and focused exclusively on final odds.

The contribution of our study is to move beyond this convention by testing whether expected returns depend not only on final odds but also on the trajectory of odds movements during the betting period. To this end, we extend the standard regression framework used to detect the FLB—where realized returns are regressed on final odds—by incorporating measures of interim odds dynamics as explanatory variables. This approach enables us to ask whether horses with identical final odds but distinct interim trajectories yield systematically different average returns, thereby providing a direct empirical test of path dependence. Exploiting the five-minute update structure of the JRA data, we construct multiple indicators of odds changes and identify the phases of the betting process in which path dependence is most pronounced.

Our main findings reveal a systematic relationship between expected returns and odds movements in the final five minutes before post time. In particular, realized returns are negatively related to last-minute changes in odds: horses experiencing a late surge in popularity—manifested as declining odds—tend to yield higher realized returns than those with identical final odds but no such movement. This pattern suggests that late-stage odds dynamics reflect the behavior of informed bettors who strategically place their wagers at the very end of the betting period in response to private signals not yet incorporated into market prices (Ottaviani and Sørensen 2006, 2009).[2]

Another key finding concerns how incorporating late-stage odds dynamics alters the conventional relationship between final odds and realized returns. Once these dynamics are taken into account, the negative correlation typically interpreted as the FLB becomes markedly weaker. This attenuation indicates that final odds alone are not sufficient statistics, and that static models relying exclusively on them may yield biased parameter estimates of risk preferences or belief

[2]The systematic relationship we document between expected returns and last-minute odds movements should be interpreted as an *in-sample* correlation rather than evidence of ex-ante predictability or arbitrage opportunities. Bettors cannot revise their actions based on final odds, nor can they submit limit orders contingent on them. Indeed, when we consider forecasts based only on odds information available up to five minutes before post time, we find that the level of odds at that point is negatively related to realized returns—similar to the well-known correlation between final odds and returns—whereas earlier odds changes are largely unrelated to returns, with only minor exceptions. Hence, while the FLB attributable to risk preferences or probability perceptions remains present, *ex-ante predictability based on interim odds changes is very limited.*

distributions.

These findings carry two key implications. First, we call into question the validity of static empirical strategies for inferring market expectations or bettor preferences, as our results show that expected returns depend not only on the level of final odds but also on last-minute odds movements. Second, we provide empirical support for the information-based explanation of the FLB developed by Ottaviani and Sørensen (2009), who demonstrate that the bias can also arise when last-minute betting occurs simultaneously and prevents the full incorporation of private information. In such settings, final odds cannot fully aggregate information and contain a "surprise" component to which bettors cannot respond, generating return differentials that static pricing models fail to capture.

Related Literature A wide range of mechanisms has been proposed to explain the FLB, which can be broadly grouped into two categories. The first attributes the bias to risk preferences under rational expectations, emphasizing either risk-loving behavior (Weitzman 1965) or systematic risk misperceptions (Jullien and Salanié 2000, Bradley 2003, Snowberg and Wolfers 2010). A recent contribution by Chiappori et al. (2019) develops and estimates a comprehensive structural model that incorporates heterogeneity in risk preferences, accommodating both expected utility and non-expected utility representations. The second category focuses on heterogeneity in information under the assumption of risk neutrality, attributing the FLB to differences in bettors' subjective beliefs. For instance, Ali (1977) shows that heterogeneous beliefs alone can generate the FLB even when all agents are risk-neutral, and Gandhi and Serrano-Padial (2015) employ a demand-estimation approach to document the coexistence of informed bettors and noise traders. While these studies provide valuable insights, they are grounded in static frameworks that implicitly assume bettors can condition their decisions on final odds.

A notable exception—and the study most closely related to ours—is Ottaviani and Sørensen (2009). They analyze parimutuel betting markets with homogeneous bettors who possess imperfect information and can condition their actions on private signals but not on final odds. Their key insight is that in the Bayesian-Nash equilibrium with last-minute simultaneous betting that does not condition on final odds, the surprise revealed by the final odds remains unincorporated into bets. As a result, final odds fail to fully aggregate information, and the FLB arises.[3] Ottaviani and Sørensen (2010) extend this framework and propose *indirect* empirical tests that rely solely on final odds. By contrast, our study provides more direct evidence on the last-minute surprise hypothesis by exploiting rich intra-race data on odds trajectories observed throughout the betting period.

In addition, several studies emphasize how institutional features of betting markets shape their informational efficiency. Smith, Paton and Williams (2006) show that market efficiency depends on

[3] For an overview of information-based explanations, see Bergemann and Ottaviani (2021).

the institutional setting by analyzing Betfair online betting in UK horse racing, where bettors post and accept wagers in a manner that effectively allows contingent price offers, in sharp contrast to parimutuel systems. Importantly, they find that the FLB, well documented in bookmaker markets, is much weaker in this exchange environment. Subsequent work by Franck, Verbeek and Nüesch (2010) provides further evidence from European football betting markets, showing that the mechanisms by which odds are set and adjusted affect the degree of efficiency. These studies highlight that the efficiency of betting markets depends not only on bettors' preferences or beliefs, but also on the institutional design governing how odds are determined.

Structure of the Study The remainder of the study is organized as follows. Section 2 introduces the dataset and documents key empirical facts. Section 3 sets out the econometric framework. Section 4 reports the main findings, while Section 5 provides additional analyses and discussion. Finally, Section 6 concludes.

2 Data

This section begins by outlining the institutional background of horse racing in Japan, and then describes the dataset used in our empirical analysis. We conclude by presenting several key empirical patterns that motivate our investigation and guide the subsequent econometric analysis.

2.1 Institutional Details

In Japan, horse racing is organized by the JRA and the National Association of Racing (NAR). The JRA oversees racing events at ten major racecourses (e.g., Tokyo, Nakayama, Kyoto, and Hanshin) in metropolitan areas, while the NAR is responsible for local races held throughout the country. Races at the ten major tracks administered by the JRA are referred to as *Chuo Keiba* (meaning "central horse racing"), comprising roughly 3,400 races annually. As one of the largest and most profitable racing organizations worldwide, the JRA maintains a fully integrated infrastructure and, in recent years, has reported annual betting turnover exceeding 2.8 trillion yen (approximately 19 billion USD at an exchange rate of 145 yen to the dollar as of 2024).[4]

This study focuses exclusively on JRA races, for which past race records are systematically archived in a uniform format. All JRA races are conducted under a parimutuel betting system, in which odds are dynamically updated according to the distribution of wagers and final payouts are determined by the closing odds. Races are held primarily on weekends and Japanese national holidays. Betting for Saturday races opens at 6:30 p.m. on the preceding Friday, and for Sunday

[4]Beyond race administration, the JRA also manages training centers, equine development facilities, and a professional school for jockeys and stable personnel.

races at 7:30 p.m. on the preceding Saturday. Wagering closes roughly one minute before post time (see Supplementary Appendix A.1 for further details), and tickets can be purchased in units of 100 JPY (approximately 0.69 USD).

Our analysis focuses on the win (*tanshō*) pool, where the JRA deducts 20% from the total pool to cover taxes and operating expenses. Odds—defined as the gross payout per 1 JPY bet—are given by

$$R_i = \frac{(1-0.2)\sum_{i\in I} W_i}{W_i} = \frac{0.8}{s_i}, \tag{1}$$

where W_i denotes the total amount wagered on horse i, I is the set of competing horses, and $s_i \equiv W_i / \sum_{i\in I} W_i$ represents the share of total bets placed on horse i.[5] By construction, odds are lower for horses that attract more wagers (favorites) and higher for those that attract fewer (longshots).

It should be noted that equation (1) applies to both final and interim odds. Interim odds are calculated from the cumulative wagers placed up to a given point in time and thus provide an intermediate aggregation of bettors' actions during the betting period. While interim odds do not determine payouts, they offer valuable information on the real-time evolution of market sentiment.

The JRA also operates a supplemental payout scheme known as "JRA Plus 10." If the payout for a winning 100-JPY ticket would otherwise be exactly 100 JPY (i.e., odds of 1.0), it is increased to 110 JPY—unless the bonus would exceed the pool surplus, in which case the payout reverts to 100 JPY. Accordingly, for horses with final odds of 1.0, the realized odds may be either 1.0 or 1.1 depending on race-specific conditions. To avoid ambiguity in return calculations, we exclude from our sample any race in which at least one horse had final odds of exactly 1.0. Such cases are extremely rare, representing only 0.0007% of horses in our 20-year dataset.

Finally, if a horse is withdrawn or declared a non-starter after betting has commenced, all tickets involving that horse are refunded. However, posted odds are not adjusted in real time to reflect these changes, which may generate discrepancies between observed odds and actual market expectations. To ensure data integrity, we exclude from our analysis all races in which any horse was withdrawn after betting had opened.

2.2 Horse-Race-Time-Level Panel Dataset

Our empirical analysis relies on data from the JRA-VAN database, a licensed service that provides structured records from the JRA. For each race, the dataset contains rich metadata including race-track, date and time, race classification, track condition, course distance, and horse-level attributes (e.g., name, age, sex, and jockey). It also reports both final and interim win odds, race outcomes

[5]In Japan, odds in horse racing are quoted as gross payouts inclusive of the original stake. For example, odds of 2.5 imply that a 1 JPY bet yields 2.5 JPY if successful. This corresponds to fractional odds of 3/2, yielding 1.5 JPY in net profit. In general, JRA odds R_i correspond to fractional odds of $R_i - 1$.

(finishing positions and official times), and aggregate betting volumes. Descriptive statistics for the sample are presented in Supplementary Appendix A.1.

A distinctive feature of the JRA-VAN platform is the availability of high-frequency interim odds. In addition to the final odds published at the close of wagering, the JRA releases interim odds at roughly five-minute intervals throughout the betting window, until one minute before post time. These updates make it possible to trace the evolution of market expectations in real time.

Interim odds are obtained via automated polling of the JRA's official servers. Although betting occurs continuously, the precise timing of odds updates can vary slightly across races. To ensure comparability and temporal consistency, we harmonize the data into fixed five-minute intervals. This yields a balanced horse-race-time-level panel dataset that captures the dynamics of odds formation over the entire betting period. The structure of this dataset enables us to test whether movements in interim odds provide incremental predictive content for realized returns, conditional on final odds.

2.3 Basic Facts

We begin by documenting several key empirical patterns regarding the dynamics of odds and their relationship with realized returns.

Realized Returns and Final Odds Figure 1 illustrates the relationship between realized returns and final odds. Horses are grouped into bins based on their odds using fixed-width intervals (e.g., $[1, 11), [11, 21), \ldots$). For each group $g \in \mathcal{G}$, we compute the average realized *gross return* as

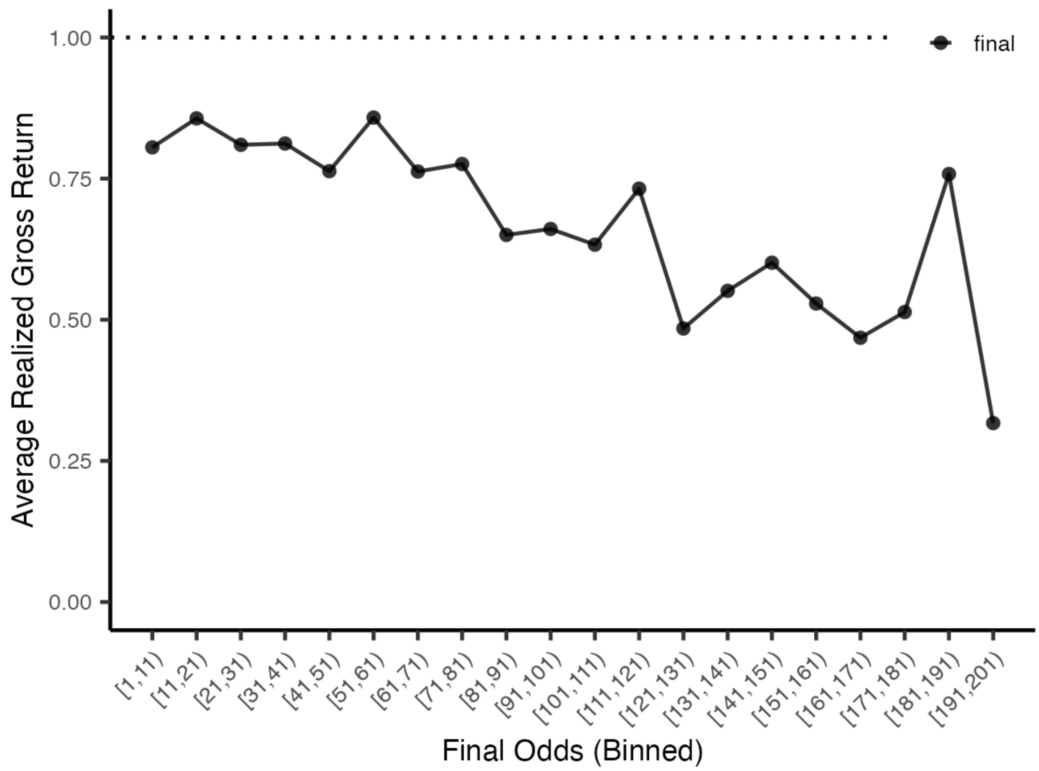
$$\frac{1}{|I_g|} \sum_{i \in I_g} \mathbb{1}_{\{win_i = 1\}} R_i^*,$$

where I_g denotes the set of horses in group g, and R_i^* represents the final odds for horse i, inclusive of the original stake. Losing horses yield zero returns and therefore do not contribute to the group average.

The figure reveals a clear negative relationship between final odds and realized returns: horses with lower final odds (favorites) yield higher average returns, whereas those with higher odds (longshots) generate lower returns. This pattern is consistent with the well-established FLB, whereby bets on longshots systematically underperform those on favorites.

Temporal Distribution of Bets Figure 2 shows the temporal distribution of wagering activity across races. For each race, the cumulative amount wagered by post time is normalized to one, and the figure plots the proportion of total bets placed at each point prior to post time. The data reveal a pronounced acceleration in betting as post time nears: only about 25% of wagers

Figure 1: Odds versus Returns



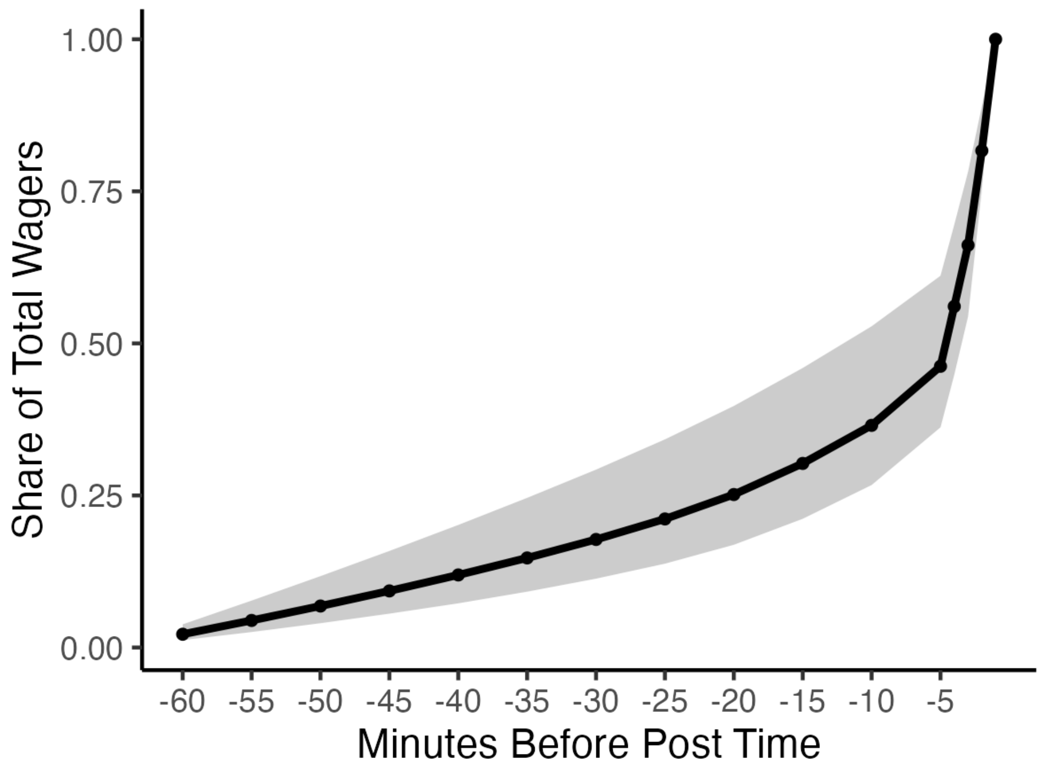
Notes: The solid black line plots the average realized gross return for each odds group $g \in \mathcal{G}$, calculated as $\frac{1}{|I_g|}\sum_{i \in I_g} \mathbb{1}_{\{\text{win}_i=1\}} R_i^*$, where R_i^* denotes the final odds for horse i. Odds are grouped into intervals of width 10, and bins with fewer than 1,000 observations are excluded, following Jullien and Salanié (2000).

are placed by 20 minutes before post time, whereas nearly half are submitted within the final five minutes. This pattern underscores the substantial clustering of wagers in the closing moments of the betting window.[6]

Odds Changes and Outcomes We next examine how interim odds dynamics relate to the association between final odds and realized returns. Figure 3 presents a series of comparisons based on whether a horse's odds increased or decreased during specific time intervals prior to post time. Panel (a) focuses on the final five-minute window and classifies horses according to whether their odds rose or fell during this period. The results show that, conditional on similar final odds, horses whose odds declined—suggesting a late surge in betting interest—consistently yield higher

[6]This concentration of late wagering parallels the "sniping" behavior observed in online auctions such as eBay, where participants strategically delay bids until the final moments (Roth and Ockenfels 2002, Ockenfels and Roth 2006, Bajari and Hortacsu 2004). In the context of parimutuel betting, Ottaviani and Sørensen (2006) develop a theoretical framework in which informed bettors postpone their wagers to avoid revealing private information to the market. Such strategic delay results in a clustering of bets immediately before post time.

Figure 2: Temporal Distribution of Wagering Activity Before Post Time



Notes: The figure plots the share of cumulative wagers placed at five-minute intervals from 60 minutes to 5 minutes before post time, and at one-minute intervals during the final five minutes. For each race, total betting volume by post time is normalized to one. The solid black line indicates the median share at each time point, and the shaded region denotes the 95% interval across races.

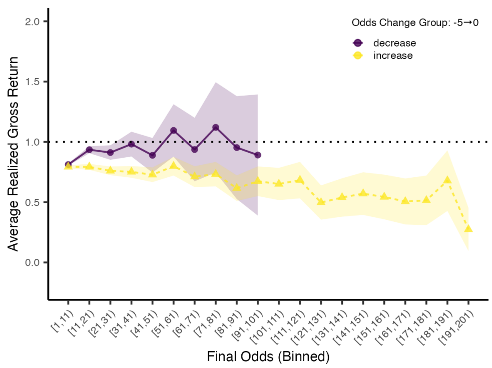
average realized returns than those whose odds increased.

Panels (b)–(f) extend the analysis to earlier five-minute windows: 5 to 10, 10 to 15, 15 to 20, 20 to 25, and 25 to 30 minutes prior to post time. Unlike the final interval, odds movements in these earlier periods display no systematic association with subsequent returns, suggesting that early fluctuations carry little predictive content.

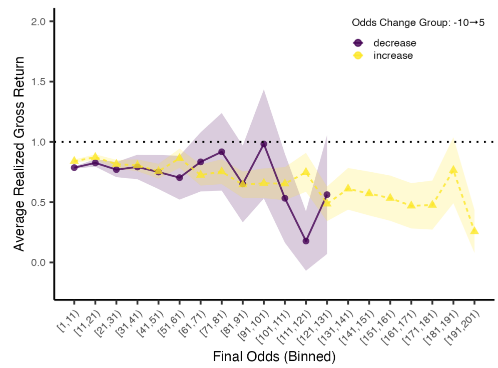
Taken together, these findings indicate that interim odds dynamics are informative about expected returns only in the final stages of the betting period. This pattern underscores the distinct role of late-stage wagering activity in shaping final prices and highlights the importance of accounting for path dependence in empirical analyses of betting markets. In the next section, we formalize these insights within an econometric framework to quantify the magnitude of these dynamic relationships.

Figure 3: Expected Return versus Interim Odds Changes

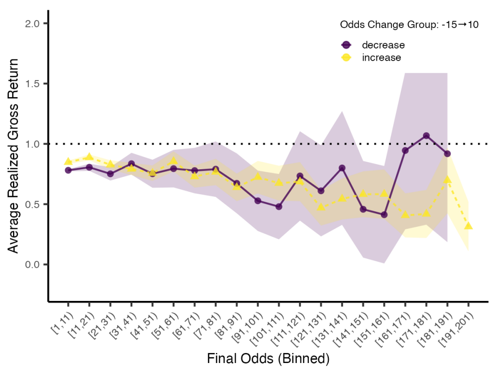
(a) -5 to 0 (final)



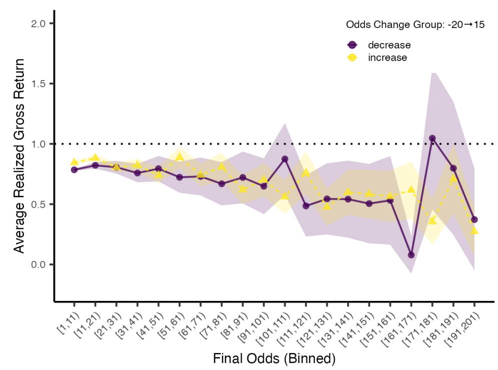
(b) -10 to -5



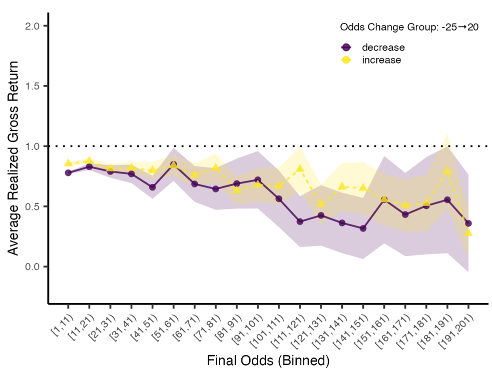
(c) -15 to -10



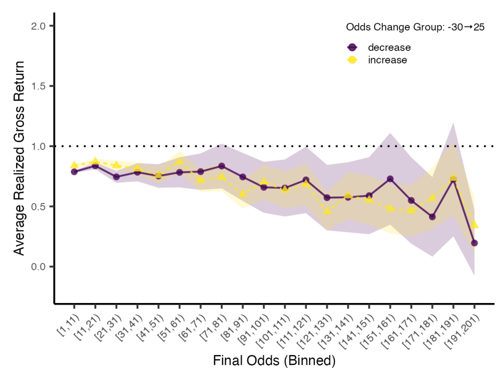
(d) -20 to -15



(e) -25 to -20



(f) -30 to -25



Note: The horizontal axis represents final odds, grouped into bins of width 10. Panels (a)–(f) compare average realized returns for horses whose odds increased versus decreased during specified five-minute intervals prior to post time. Panel (a) covers the final interval (0 to 5 minutes before post time), while Panels (b)–(f) examine earlier intervals: 5 to 10, 10 to 15, 15 to 20, 20 to 25, 25 to 30 minutes before post time, respectively. Within each panel, the two lines plot average realized returns for horses with increasing and decreasing odds during the corresponding interval. Shaded areas denote 95% confidence intervals. Following Jullien and Salanié (2000), bins are excluded if either subgroup contains fewer than 1,000 observations.

3 Econometric Strategy

The preceding section documented that odds movements in the final five minutes before post time—the window in which nearly half of all bets are placed—are systematically related to race outcomes. This section develops a formal econometric framework to quantify the presence and strength of such path dependence in expected returns.

We begin with the baseline regression model:

$$\mathbb{1}_{\{win_i=1\}} R_i^* = \alpha + \beta R_i^* + \delta \, \mathrm{OddsChange}_i + \gamma \, R_i^* \times \mathrm{OddsChange}_i + \varepsilon_i,$$

where R_i^* denotes the final odds for horse i; $\mathrm{OddsChange}_i$ is a variable (or vector of variables) summarizing the trajectory of interim odds; ε_i is a mean-zero idiosyncratic error term. The dependent variable equals the gross return to a 1 JPY bet on horse i: R_i^* if the horse wins and zero otherwise. The conditional expectation is therefore

$$\mathbb{E}[\mathbb{1}_{\{win_i=1\}} R_i^* \mid R_i^*, \mathrm{OddsChange}_i] = \alpha + \beta R_i^* + \delta \, \mathrm{OddsChange}_i + \gamma \, R_i^* \times \mathrm{OddsChange}_i,$$

where $\mathbb{E}[\cdot \mid X]$ denotes the expectation conditional on covariates X.

This specification highlights the three parameters of interest. The intercept, α, represents the baseline expected return in the absence of variation in odds, while β traces how expected returns vary with the level of final odds, thus testing for the conventional FLB. The coefficient δ isolates the correlation between expected returns and the odds trajectory, conditional on the final odds level, thereby capturing whether interim odds contain predictive information beyond final odds. Finally, γ examines whether this correlation itself varies systematically with the magnitude of final odds, providing a test for heterogeneity in path dependence across the odds distribution.

Under the null hypothesis of full-information rational expectations and risk neutrality, final odds fully incorporate all available information, rendering bettors indifferent across horses. Accordingly, neither the level nor the trajectory of odds should systematically affect expected returns. Formally, this null corresponds to $\beta = \delta = \gamma = 0$, with the intercept $\alpha = 0.8$ capturing the 20% takeout rate imposed under the JRA parimutuel system. Thus, the expected gross return from a 1 JPY bet is uniformly 0.8, independent of both the magnitude and the evolution of odds.

4 Results

Table 1 reports the coefficient estimates from the regression model. Column (1) shows the results from a restricted specification that excludes path dependence ($\delta = \gamma = 0$) and includes only final odds as a regressor. In this model, the coefficient β captures the relationship between final odds and expected returns. The estimate, $\beta = -0.0016$, is statistically significant at the 1% level,

Table 1: Estimation Results: Winning Odds

	(1)	(2)	(3)	(4)	(5)
Constant	0.8338	0.8452	0.8443	0.8435	0.8423
	(0.0074)	(0.0080)	(0.0082)	(0.0082)	(0.0082)
R_i^*	-0.0016	-0.0012	-0.0012	-0.0013	-0.0014
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
$\frac{\Delta R_{i,[-5,0]}}{R_{i,-5}}$		-0.3559			
		(0.0381)			
$\frac{\Delta R_{i,[-10,0]}}{R_{i,-10}}$			-0.1674		
			(0.0250)		
$\frac{\Delta R_{i,[-15,0]}}{R_{i,-15}}$				-0.1052	
				(0.0210)	
$\frac{\Delta R_{i,[-20,0]}}{R_{i,-20}}$					-0.0702
					(0.0176)
$R_i^* \times \frac{\Delta R_{i,[-5,0]}}{R_{i,-5}}$		0.0002			
		(0.0003)			
$R_i^* \times \frac{\Delta R_{i,[-10,0]}}{R_{i,-10}}$			0.0000		
			(0.0002)		
$R_i^* \times \frac{\Delta R_{i,[-15,0]}}{R_{i,-15}}$				0.0001	
				(0.0001)	
$R_i^* \times \frac{\Delta R_{i,[-20,0]}}{R_{i,-20}}$					0.0001
					(0.0001)
Num.Obs.	894127	894127	894127	894127	894127
R2	0.001	0.001	0.001	0.001	0.001
R2 Adj.	0.001	0.001	0.001	0.001	0.001

Notes: This table reports OLS estimates of the coefficients from the regression:

$$\mathbb{1}_{\{win_i=1\}} R_i^* = \alpha + \beta R_i^* + \delta \, \text{OddsChange}_i + \gamma \, R_i^* \times \text{OddsChange}_i + \varepsilon_i.$$

The variable $\Delta R_{i,[-\tau,0]}/R_{i,-\tau} \equiv (R_i^* - R_{i,-\tau})/R_{i,-\tau}$ represents the rate of change in odds over the final τ minutes before post time.

indicating a robust negative association. Economically, a one-unit increase in final odds reduces expected gross returns by 0.0016 JPY per 1 JPY bet, equivalent to 0.16 JPY for a standard 100 JPY wager—the minimum betting unit in the JRA system. This result rejects the benchmark of risk-neutral rational expectations and provides evidence consistent with the FLB.

Columns (2)–(5) introduce alternative specifications that incorporate measures of odds trajectories to test for path dependence. These models include δ and γ as additional parameters, allowing expected returns to vary not only with final odds but also with how those odds evolved.

Column (2) focuses on the rate of odds change during the final five minutes before post time, defined as $\Delta R_{i,[-5,0]}/R_{i,-5} \equiv (R_i^* - R_{i,-5})/R_{i,-5}$, where $R_{i,-\tau}$ denotes the interim odds for horse i observed $\tau > 0$ minutes before post time. The estimated regression is

$$\mathbb{1}_{\{win_i=1\}} R_i^* = \alpha + \beta R_i^* + \delta \, \frac{\Delta R_{i,[-5,0]}}{R_{i,-5}} + \gamma \, R_i^* \times \frac{\Delta R_{i,[-5,0]}}{R_{i,-5}} + \varepsilon_i.$$

The estimates from Column (2) show a statistically significant association between late-stage odds movements and expected returns. The coefficient on the odds-change term is negative and significant ($\delta = -0.3559$), indicating that horses whose odds decline in the final five minutes tend to yield systematically higher expected returns, conditional on their final odds. By contrast, the interaction term is small and statistically insignificant ($\gamma = 0.0002$), suggesting that this relationship does not vary systematically with the level of final odds.

Notably, once late-stage odds movements are accounted for, the coefficient on final odds decreases in magnitude from $\beta = -0.0016$ in Column (1) to $\beta = -0.0012$ in Column (2) while remaining statistically significant. This attenuation implies that part of the negative association previously attributed to final odds is captured by the dynamics of late betting, underscoring the predictive content of interim odds beyond their final level.[7]

Columns (3)–(5) extend the analysis by considering longer windows prior to post time: 10, 15, and 20 minutes, respectively. For each window, the odds change rate is defined as $\Delta R_{i,[-\tau,0]}/R_{i,-\tau} \equiv (R_i^* - R_{i,-\tau})/R_{i,-\tau}$, where $R_{i,-\tau}$ denotes the interim odds observed τ minutes before post time. The estimated coefficients on these odds-change variables are smaller in magnitude for longer windows, suggesting that odds movements occurring closer to post time carry stronger predictive content for expected returns, whereas earlier movements are less systematically related to outcomes.

In summary, the regression results demonstrate that expected returns are systematically related not only to the level of final odds but also to the trajectory of interim odds, particularly in the final minutes before post time. While the conventional specification in Column (1) detects the favorite-longshot bias, incorporating measures of late-stage odds movements in Columns (2)–(5) reveals additional predictive content that attenuates the role of final odds.[8] These findings provide direct evidence of path dependence in betting markets, indicating that final odds alone do not fully capture the information embedded in wagering activity. These results are consistent with the information-based explanation of Ottaviani and Sørensen (2009), which attributes the FLB to the

[7] These results should not be interpreted as evidence of *ex-ante* return predictability. The correlations between odds changes and realized returns reflect only *ex-post* associations, because bettors cannot react after observing the final odds nor place contingent orders based on them. A more detailed discussion of the potential for *ex-ante* return predictability is provided in Supplementary Appendix A.2.

[8] Our results are robust to the inclusion of race fixed effects. See the robustness check in Supplementary Appendix A.4.

behavior of informed bettors who strategically place their wagers at the very end of the betting period in response to private signals. In this setting, final odds embed a "surprise" component that generates predictable return differentials, even under risk neutrality.

5 Discussion

The preceding section provided empirical evidence consistent with the information-based explanation of the FLB proposed by Ottaviani and Sørensen (2009), using interim odds data to uncover dynamic patterns in betting behavior. In this section, we further assess the validity of this account by testing additional theoretical predictions derived from the generalized model of Ottaviani and Sørensen (2010). Their framework extends the stylized setting of Ottaviani and Sørensen (2009) to incorporate broader aspects of market structure. By examining whether these extended predictions are borne out in our data, we evaluate the robustness and explanatory scope of the information-based interpretation of the FLB.

Racetrack Conditions Within the information-based framework, the magnitude of the FLB depends in part on how precisely private information predicts race outcomes. Ottaviani and Sørensen (2010) emphasize that environmental conditions—particularly racetrack quality—affect this precision. Under adverse conditions such as heavy rain or track degradation, increased uncertainty reduces the informativeness of private signals. Consequently, the FLB is expected to weaken when track conditions deteriorate, as bettors' ability to exploit informational advantages becomes more limited.

To test this hypothesis, we incorporate measures of racetrack condition—reported in the JRA data on a four-level ordinal scale (*Very Good*, *Moderate*, *Poor*, *Very Bad*)—into our regression framework. We pursue two empirical strategies.

First, we introduce an interaction between final odds and a dummy variable for *Very Bad* conditions in the expected return regression:

$$\mathbb{1}_{\{win_i=1\}} R_i^* = \alpha + \beta R_i^* + \beta_{condition}\, R_i^* \times \mathbb{1}_{\{condition=Very\ Bad\}} + \varepsilon_i.$$

As reported in Column (1) of Table 2, the coefficient $\beta_{condition}$ is positive and statistically significant, implying that the negative association between final odds and realized returns is significantly weaker under poor track conditions. This result is consistent with the theoretical prediction that adverse environments dilute the informational advantage of bettors.

Second, we replace the binary variable with an ordinal measure $Z \in \{1, 2, 3, 4\}$, corresponding to *Very Good*, *Moderate*, *Poor*, and *Very Bad* conditions, respectively. This specification allows us to test whether the FLB declines progressively as track quality worsens. Column (2) of Table 2

Table 2: Additional Tests of the Information-Based Model's Predictions

	(1)	(2)	(3)	(4)
R_i^*	-0.001712	-0.001955	-0.001580	-0.000110
	(0.000102)	(0.000179)	(0.000100)	(0.000008)
$R_i^* \times 1$(Racetrack Condition = Very Bad)	0.000504			
	(0.000221)			
$R_i^* \times$(Racetrack Condition Poorness)		0.000266		
		(0.000125)		
$R_i^* \times$(Num. of Wagers)			0.000001	-0.000001
			(0.000080)	(0.000001)
Odds Data	W	W	W	Q
Num.Obs.	894127	894127	893824	1315425
R2	0.001	0.001	0.001	0.000
R2 Adj.	0.001	0.001	0.001	0.000

Notes: W = winning odds, Q = quinella odds. The number of wagers is computed from the number of tickets sold. Racetrack condition is categorized into four levels by increasing surface moisture: 1 = Very Good, 2 = Moderate, 3 = Poor, 4 = Very Bad. Column (2) includes the race track condition levels as numerical variables.

confirms that the bias becomes less pronounced with deteriorating conditions, in line with model predictions.

Bettor Population Size Ottaviani and Sørensen (2010) predict that the number of participating bettors—particularly informed ones—affects the magnitude of the FLB. As the number of informed bettors increases, so does the aggregate volume of private information. If this information is not fully incorporated due to the simultaneity of last-minute betting, the FLB should become more pronounced.

We test this hypothesis by examining whether the magnitude of the bias varies systematically with the total amount wagered in each race, which serves as a proxy for market participation. Column (3) of Table 2 reports the estimates. Contrary to the theoretical prediction, we find no statistically significant evidence that the FLB intensifies in races with greater participation; the estimated interaction term is close to zero.

Two factors may account for this null result. First, most JRA races already attract a large number of participants, so marginal increases may add little additional informational content to the market. Second, in high-profile races such as graded stakes, higher participation often reflects the presence of casual or infrequent bettors, whose decisions may be less informed or more noise-driven than those of regular bettors. Consequently, the effective informational content of additional wagers may be diluted, weakening the link between participation size and the strength of the FLB.

Number of Possible Outcomes: Quinella Bets Finally, Ottaviani and Sørensen (2010) theorize that the magnitude of the FLB depends on the complexity of the betting instrument, particularly the number of possible outcomes. As the number of potential outcomes increases relative to the number of bets placed, the informational content embedded in final odds diminishes. Consequently, the negative association between final odds and realized returns should weaken, and in some cases may even reverse the typical FLB pattern.

Column (4) examines odds in the quinella (*umaren*) market, in which bettors select two horses and the ticket pays if both finish first and second in any order.[9] The results show that, although the negative association remains, its slope is considerably smaller than in the win market (see also Supplementary Appendix A.3), consistent with the prediction of Ottaviani and Sørensen (2010).

Summary Overall, our findings provide mixed support for the predictions of the information-based framework. We find strong evidence that unfavorable racetrack conditions attenuate the FLB and that more complex betting instruments are associated with a weaker bias. By contrast, we do not find empirical confirmation that greater numbers of bettors amplify the FLB, suggesting that market participation may not translate directly into greater informational content.

6 Conclusion

This study examined the dynamics of information aggregation in parimutuel betting markets by analyzing interim odds throughout the betting period. Conventional static betting market models assume that bettors condition their choices on final odds, and prior research has therefore focused on the static relationship between final odds and realized returns. Our evidence, however, demonstrates that odds trajectories—especially those in the final minutes before post time—play a critical role in shaping expected returns.

Our analysis shows that horses experiencing last-minute declines in win odds, indicative of rising popularity, consistently deliver higher average returns than horses with identical final odds whose odds increased over the same interval. These late-stage dynamics attenuate the negative correlation between final odds and realized returns that is conventionally interpreted as the FLB.

These findings carry several implications. First, they suggest that late-stage odds movements reflect the behavior of informed bettors who strategically place their wagers at the very end of the betting period in response to private signals. Second, they indicate that final odds may not fully aggregate information, implying that static models relying exclusively on them may yield biased parameter estimates of risk preferences or belief distributions. More broadly, our results challenge the sufficiency of final odds assumed in conventional static betting market models and

[9] Unlike exacta bets, which require the correct order of finish, quinella bets only require that the two selected horses occupy the top two positions, regardless of order.

provide complementary support for information-based explanations of the FLB (Ottaviani and Sørensen 2009, 2010). They also highlight how market design features—such as the absence of limit orders—shape information aggregation, and point to future research on whether similar path dependence arises in other market environments.

References

Ali, Mukhtar M., "Probability and utility estimates for racetrack bettors," *Journal of Political Economy*, 1977, *85* (4), 803–815.

Bajari, Patrick and Ali Hortacsu, "Economic insights from internet auctions," *Journal of Economic Literature*, 2004, *42* (2), 457–486.

Barseghyan, Levon, Francesca Molinari, Ted O'Donoghue, and Joshua C. Teitelbaum, "Estimating risk preferences in the field," *Journal of Economic Literature*, 2018, *56* (2), 501–564.

Bergemann, Dirk and Marco Ottaviani, "Information markets and nonmarkets," in "Handbook of Industrial organization," Vol. 4, Elsevier, 2021, pp. 593–672.

Bradley, Ian, "The representative bettor, bet size, and prospect theory," *Economics Letters*, 2003, *78* (3), 409–413.

Chiappori, Pierre-André, Bernard Salanié, François Salanié, and Amit Gandhi, "From aggregate betting data to individual risk preferences," *Econometrica*, 2019, *87* (1), 1–36.

Franck, Egon, Erwin Verbeek, and Stephan Nüesch, "Prediction accuracy of different market structures — bookmakers versus a betting exchange," *International Journal of Forecasting*, 2010, *26* (3), 448–459. Sports Forecasting.

Gandhi, Amit and Ricardo Serrano-Padial, "Does belief heterogeneity explain asset prices: The case of the longshot bias," *Review of Economic Studies*, 2015, *82* (1), 156–186.

Gillen, Benjamin J., Charles R. Plott, and Matthew Shum, "A pari-mutuel-like mechanism for information aggregation: A field test inside Intel," *Journal of Political Economy*, 2017, *125* (4), 1075–1099.

Griffith, Richard M., "Odds adjustments by American horse-race bettors," *American Journal of Psychology*, 1949, *62* (2), 290–294.

Hausch, Donald B. and William T. Ziemba, "Efficiency of sports and lottery betting markets," *Handbooks in Operations Research and Management Science*, 1995, *9*, 545–580.

Hurley, William and Lawrence McDonough, "A note on the Hayek hypothesis and the favorite-longshot bias in Parimutuel betting," *American Economic Review*, 1995, *85* (4), 949–55.

Jullien, Bruno and Bernard Salanié, "Estimating preferences under risk: The case of racetrack bettors," *Journal of Political Economy*, 2000, *108* (3), 503–530.

_ **and** _ , "Empirical evidence on the preferences of racetrack bettors," in "Handbook of Sports and Lottery Markets," Elsevier, 2008, p. 27.

Kajii, Atsushi and Takahiro Watanabe, "Favorite–longshot bias in pari-mutuel betting: An evolutionary explanation," *Journal of Economic Behavior & Organization*, 2017, *140*, 56–69.

Koessler, Frederic, Charles Noussair, and Anthony Ziegelmeyer, "Information aggregation and belief elicitation in experimental parimutuel betting markets," *Journal of Economic Behavior & Organization*, 2012, *83* (2), 195–208.

Ockenfels, Axel and Alvin E. Roth, "The stability of internet auction markets," *American Economic Review*, 2006, *96* (2), 99–103.

Ottaviani, Marco and Peter Norman Sørensen, "The timing of parimutuel bets," *Unpublished*, 2006.

_ **and** _ , "The favorite-longshot bias: An overview of the main explanations," in "Handbook of Sports and Lottery Markets," Elsevier, 2008, pp. 83–101.

_ **and** _ , "Surprised by the parimutuel odds?," *American Economic Review*, 2009, *99* (5), 2129–2134.

_ **and** _ , "Noise, information, and the favorite-longshot bias in parimutuel predictions," *American Economic Journal: Microeconomics*, 2010, *2* (1), 58–85.

Plott, Charles R., Jorgen Wit, and Winston C. Yang, "Parimutuel betting markets as information aggregation devices: experimental results," *Economic Theory*, 2003, *22* (2), 311–351.

Roth, Alvin E. and Axel Ockenfels, "Last-minute bidding and the rules for ending second-price auctions: Evidence from eBay and Amazon auctions on the internet," *American Economic Review*, 2002, *92* (4), 1093–1103.

Smith, Michael A., David Paton, and Leighton Vaughan Williams, "Market efficiency in person-to-person betting," *Economica*, 2006, *73* (292), 673–689.

Snowberg, Erik and Justin Wolfers, "Explaining the favorite–long shot bias: Is it risk-love or misperceptions?," *Journal of Political Economy*, 2010, *118* (4), 723–746.

Thaler, Richard H. and William T. Ziemba, "Anomalies: Parimutuel betting markets: Racetracks and lotteries," *Journal of Economic Perspectives*, 1988, *2* (2), 161–174.

Weitzman, Martin, "Utility analysis and group behavior: An empirical study," *Journal of Political Economy*, 1965, *73* (1), 18–26.

A Supplementary Appendix

A.1 Data Details

Betting Channels and User Segmentation in JRA Races Wagers on JRA races can be placed through a variety of channels, including online platforms and physical betting venues. The primary online channel is Internet Programmed Automatic Transmission (IPAT), which requires prior registration and linkage to a domestic bank account. Another online option, JRA Direct, operates via credit card and allows same-day registration and betting, though it is accessible only via desktop browsers and does not support mobile devices. In addition to these online channels, bets may also be placed in person at JRA-operated racecourses and WINS (off-track betting facilities).

Although the JRA does not publish micro-level data on individual bettors, industry reports and publicly available aggregate statistics indicate that online channels—namely IPAT and JRA Direct—collectively account for over 70% of total wagering turnover. Betting at physical locations tends to be dominated by older bettors and is often characterized by larger per-bet amounts. In contrast, online platforms attract a broader demographic and are typically associated with smaller but more frequent wagers.

The cutoff times for purchasing JRA betting tickets vary depending on the purchase method:

1. JRA IPAT: Up to 1 minute before the race starts.
2. At the racecourse or WINS (off-track betting facilities): Up to 2 minutes before the race starts.
3. JRA Direct and telephone betting: Up to 5 minutes before the race starts.

Summary Statistics Table A.1 reports summary statistics for winning odds data from 2004 to 2023 at both the horse-race level and the horse-race-time level. Panels (a) and (b) present basic characteristics of races and individual horses, respectively. On average, horses are 3.6 years old, each race features an average field size of 14.19 starters, and the mean final win odds is 66.8 with a standard deviation of 96.35, reflecting the wide dispersion in ex ante market expectations.

Objective Winning Probability and Final Odds Figure A.1 depicts the relationship between final odds and (ex-post) winning probabilities in the JRA betting market. It illustrates a clear inverse relationship between final odds and realized win probabilities: horses with lower odds (i.e., favorites) exhibit substantially higher empirical win rates, while those with higher odds (i.e., longshots) show markedly lower chances of winning.

A.2 Ex-ante Return Predictability

The results in Section 4 document a significant association between realized returns, $\mathbb{1}_{\{win_i=1\}}R_i^*$, and changes in odds from interim to final values in the last few minutes before the race. For clarity,

Table A.1: Summary Statistics of Winning Odds Data

(a) Race level

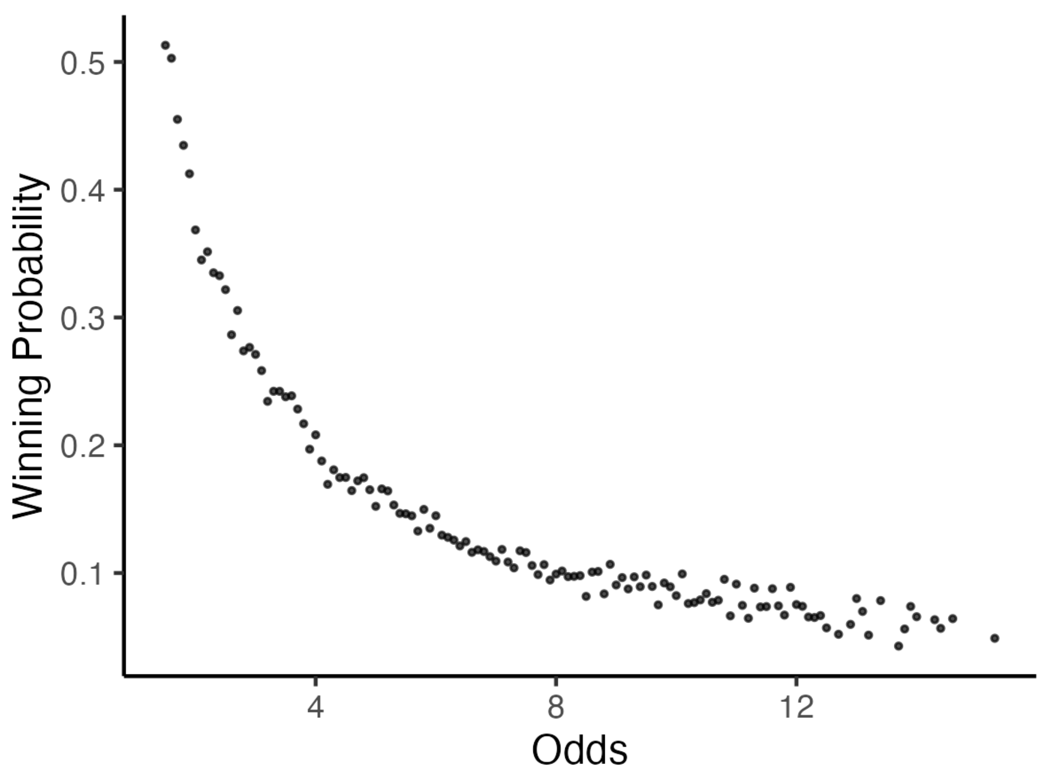
	N	Mean	SD	Min.	Max.
Num Horses	63372	14.19	2.61	4.00	18.00
Distance	63372	1663.91	444.68	1000.00	4260.00
Num. of Wagers (Mil)	63351	0.40	0.73	0.01	42.63
Racetrack Quality	63372	3.57	0.80	1.00	4.00

(b) Horse-race level

	N	Mean	SD	Min.	Max.
Horse Age	895090	3.65	1.35	2.00	13.00
Realized Rank	895090	7.81	4.43	1.00	18.00
Final Odds	895090	66.80	96.35	1.10	999.90
Realized Finish Time	895090	102.05	29.92	53.80	326.40

Note: The winning odds data cover all centrally administered races in Japan from 2004 to 2023.

Figure A.1: Winning Probability: Winning Odds



Note: The database also includes pool-level betting volumes, and horse-level and race-level detailed attributes.

these results should not be taken as evidence of *ex-ante* return predictability. Rather, it should be recognized that the correlations we document reflect only an *ex-post* association, because bettors cannot react after observing the final odds nor place contingent orders based on them.

As a complementary exercise, Table A.2 reports estimates from an alternative predictive model that relies only on information available at least τ minutes before post time. For each reference point $\tau \in \{5, 10, 15, \dots\}$, we examine whether the *rates of change in odds* over the preceding 20 minutes contain predictive power for realized returns. The 20-minute window is divided into four consecutive five-minute intervals. For example, in Column (1), where the reference point is five minutes before post time, the regressors are the rates of change in odds between 5-10, 10-15, 15-20, and 20-25 minutes before the race. Column (2) shifts the reference point to ten minutes before post time, so the regressors correspond to 10-15, 15-20, 20-25, and 25-30 minutes before the race, and so forth. In this way, each column uses odds movements observed strictly before the chosen horizon, and the coefficients indicate whether such interim dynamics help predict realized returns.

The results show that, while there exist some cases where interim odds changes are statistically significant predictors of realized returns (e.g., the odds change between 5-10 minutes before the race when evaluated at $\tau = 5$, or the odds change between 15-20 minutes before the race when evaluated at $\tau = 15$), the overall evidence of ex-ante predictability is limited. At $\tau = 5$, last-minute rushes of betting activity may already be captured in the 5-10 minute interval because wagering at off-track betting facilities closes two minutes before post time. Apart from this specific case, significant coefficients appear only sporadically across horizons and intervals, and the magnitudes are generally small.

A.3 Quinella Odds Analysis

Table A.3 reports summary statistics for quinella odds from 2019 to 2023, with the shorter sample period chosen for computational feasibility (in contrast, the win odds data in Table A.1 cover 2004-2023). Compared with the win odds, two differences stand out. First, Panel (a) shows that the total amount wagered in quinella bets is larger on average (1.16 million JPY per race) than in win bets (0.40 million JPY). Yet, because the quinella market involves far more betting objects—around 92 horse pairs per race compared with about 14 horses—the amount wagered per pair is much smaller than the amount wagered per horse in win betting. Panel (b) shows that quinella odds are substantially larger, averaging over 500 for horse-pair-time observations, whereas final win odds average only about 67.

Figure A.2 presents an analogous relationship between quinella odds and realized returns. Unlike the monotonic pattern observed in win odds in Figure 1, the relationship for quinella bets appears less systematic, suggesting that the classic FLB is more muted in exotic bets such as quinella.

Figure A.3 examines whether changes in interim quinella odds affect realized returns, mirroring the analysis for win bets in Figure 3. Across all panels, we compare average returns conditional on final odds for bets whose interim odds either increased or decreased over successive five-minute intervals before post time.

Table A.2: Ex-ante Return Predictability Test

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\frac{\Delta R_{i,[-10,-5]}}{R_{i,-10}}$	-0.39089							-0.39135
	(0.04370)							(0.04384)
$\frac{\Delta R_{i,[-15,-10]}}{R_{i,-15}}$	0.00396	-0.12663						0.00372
	(0.06727)	(0.06571)						(0.06734)
$\frac{\Delta R_{i,[-20,-15]}}{R_{i,-20}}$	-0.10140	-0.20276	-0.22235					-0.10168
	(0.06103)	(0.06002)	(0.05916)					(0.06110)
$\frac{\Delta R_{i,[-25,-20]}}{R_{i,-25}}$	0.14328	0.07026	0.05684	0.01032				0.14271
	(0.06859)	(0.06827)	(0.06793)	(0.06674)				(0.06880)
$\frac{\Delta R_{i,[-30,-25]}}{R_{i,-30}}$		-0.04467	-0.05235	-0.07030	-0.06952			-0.00052
		(0.08625)	(0.08626)	(0.08612)	(0.08576)			(0.08654)
$\frac{\Delta R_{i,[-35,-30]}}{R_{i,-35}}$			-0.00658	-0.01197	-0.01234	-0.01760		0.02524
			(0.09943)	(0.09945)	(0.09942)	(0.09939)		(0.09967)
$\frac{\Delta R_{i,[-40,-35]}}{R_{i,-40}}$				-0.11015	-0.11091	-0.11071	-0.11129	-0.07098
				(0.10646)	(0.10651)	(0.10678)	(0.10673)	(0.10692)
$\frac{\Delta R_{i,[-45,-40]}}{R_{i,-45}}$					0.04237	0.04020	0.03973	0.07338
					(0.10675)	(0.10693)	(0.10690)	(0.10700)
$\frac{\Delta R_{i,[-50,-45]}}{R_{i,-50}}$						0.04449	0.04425	0.05731
						(0.10100)	(0.10099)	(0.10101)
Num.Obs.	894127	894127	894127	894127	894124	894053	894053	894053
R2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
R2 Adj.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Notes: The regressions test whether the rates of change in odds over the 20 minutes preceding a given horizon contain predictive power for realized returns. Each column corresponds to a reference horizon $\tau \in \{5, 10, 15, \dots\}$ minutes before post time, with regressors corresponding to four consecutive five-minute intervals. All regressions include a constant term, which is not reported in the table.

Table A.3: Summary Statistics of Quinella Odds Data

(a) Race level

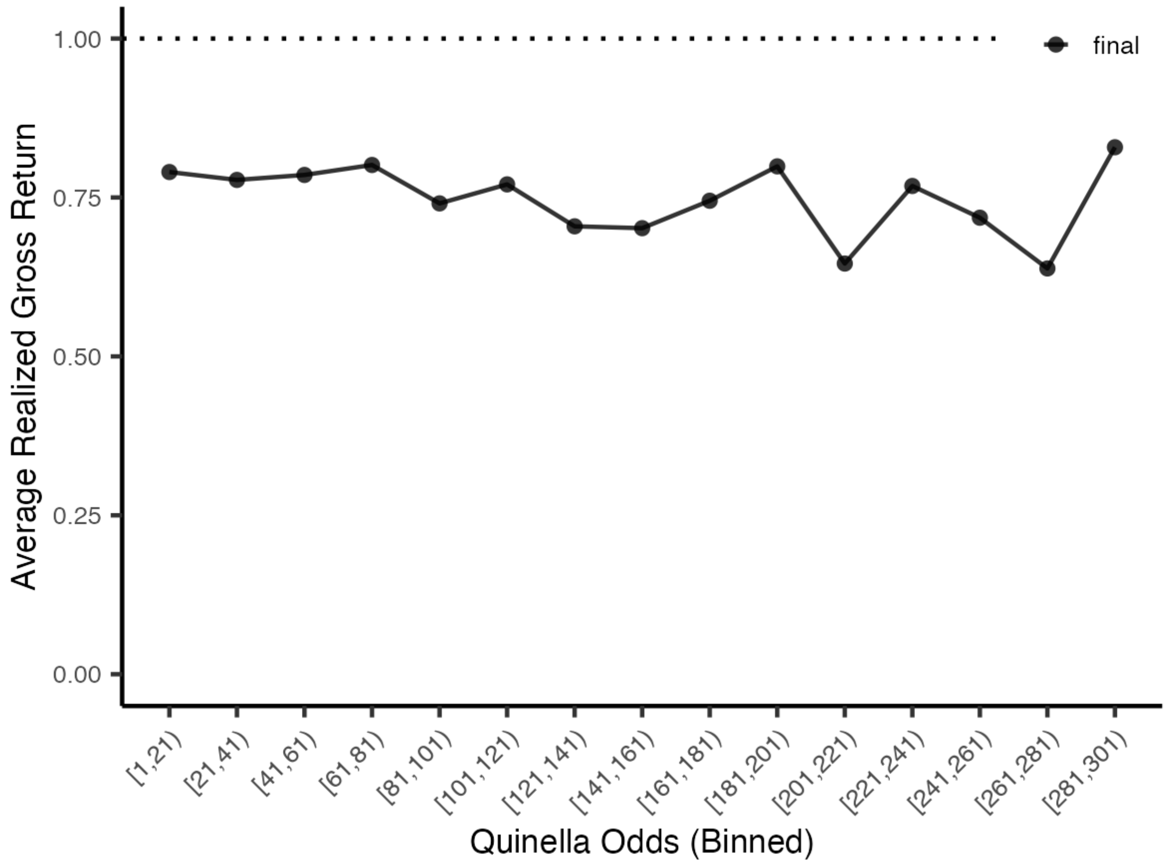
	N	Mean	SD	Min.	Max.
Num Horse Pairs	14289	92.06	34.72	10.00	153.00
Num. of Wagers (Mil)	14289	1.16	2.31	0.14	67.37

(b) Horse-pair-race level

	N	Mean	SD	Min.	Max.
Final Odds	1315425	513.33	791.98	1.10	74093.80

Note: The quinella odds data cover all centrally administered races in Japan from 2019 to 2023.

Figure A.2: Quinella Odds versus Realized Returns



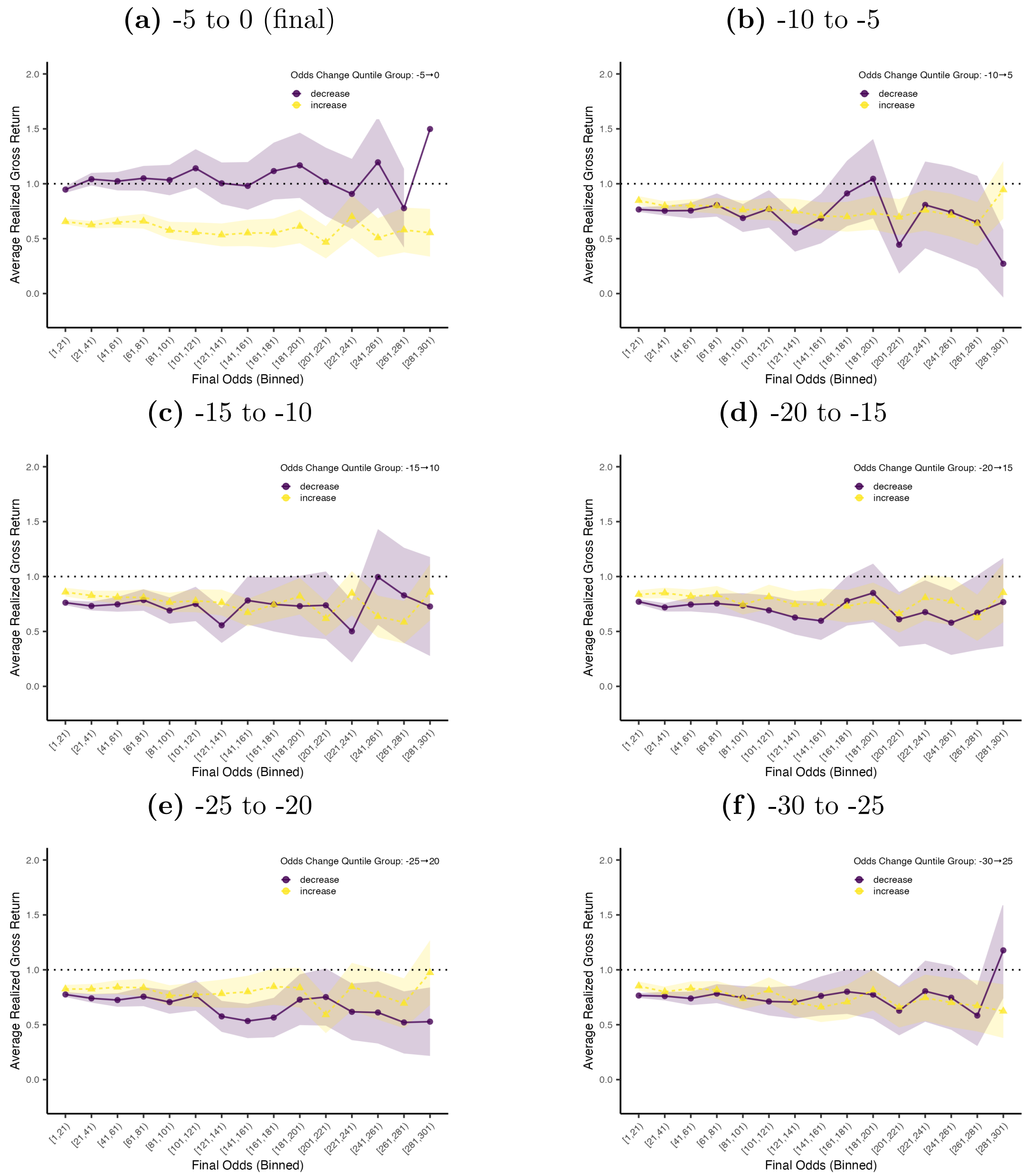
Notes: The average realized gross return for each group $g \in \mathcal{G}$ on the vertical axis is given by $\frac{1}{|I_g|}\sum_{i \in I_g} \mathbb{1}_{\{win_i=1\}} R_i^*$, where I_g denotes the set of horses in group g, and R_i^* represents the final odds for horse i. Odds are binned in intervals of width 20, and bins with fewer than 1,000 observations are excluded from the plot, as in Jullien and Salanié (2000).

As in the win bet case, we observe a divergence between the "increase" and "decrease" groups, with odds declines associated with higher realized returns and odds increases with lower returns. This pattern becomes clearly visible only in the final five minutes (Panel (a)), consistent with the pattern observed in win odds. Overall, while the classic FLB appears more muted in quinella bets, such bets nevertheless exhibit a similar form of late-stage predictability as win markets.

A.4 Robustness Analysis: Race Fixed Effects

We assess the robustness of our baseline regression results by incorporating race fixed effects and clustering standard errors at the race level. The inclusion of race fixed effects accounts for unobserved race-specific heterogeneity that may affect realized returns. A key motivation for this specification arises from the institutional detail that posted odds in JRA races are rounded to the first decimal place. As a result, the implied takeout rate—computed from observed odds—may deviate slightly from the official 20% due to rounding error. These small but systematic

Figure A.3: Expected Return vs Interim Quinella Odds Changes



Note: The horizontal axis represents final quinella odds, grouped into bins of width 10. Panels (a)–(f) compare average realized returns for horses whose odds increased versus decreased during specified five-minute intervals prior to post time. Panel (a) covers the final interval (0 to 5 minutes before post time), while Panels (b)–(f) examine earlier intervals: 5 to 10, 10 to 15, 15 to 20, 20 to 25, 25 to 30 minutes before post time, respectively. Within each panel, the two lines plot average realized returns for horses with increasing and decreasing odds during the corresponding interval. Shaded areas denote 95% confidence intervals. Following Jullien and Salanié (2000), bins are excluded if either subgroup contains fewer than 1,000 observations.

discrepancies introduce mechanical variation in expected returns across races.

Furthermore, we cluster standard errors at the race level to allow for arbitrary correlation

Table A.4: Estimation Results: Quinella Odds

	(1)	(2)	(3)	(4)	(5)
Constant	0.69829	0.77976	0.80470	0.80601	0.80031
	(0.01584)	(0.01711)	(0.01778)	(0.01802)	(0.01820)
R_i^*	-0.00011	-0.00012	-0.00013	-0.00012	-0.00012
	(0.00001)	(0.00002)	(0.00002)	(0.00002)	(0.00002)
$\frac{\Delta R_{i,[-5,0]}}{R_{i,-5}}$		-1.40233			
		(0.08494)			
$\frac{\Delta R_{i,[-10,0]}}{R_{i,-10}}$			-0.97162		
			(0.06676)		
$\frac{\Delta R_{i,[-15,0]}}{R_{i,-15}}$				-0.79996	
				(0.05799)	
$\frac{\Delta R_{i,[-20,0]}}{R_{i,-20}}$					-0.59334
					(0.04879)
$R_i^* \times \frac{\Delta R_{i,[-5,0]}}{R_{i,-5}}$		0.00023			
		(0.00005)			
$R_i^* \times \frac{\Delta R_{i,[-10,0]}}{R_{i,-10}}$			0.00018		
			(0.00003)		
$R_i^* \times \frac{\Delta R_{i,[-15,0]}}{R_{i,-15}}$				0.00013	
				(0.00002)	
$R_i^* \times \frac{\Delta R_{i,[-20,0]}}{R_{i,-20}}$					0.00010
					(0.00002)
Num.Obs.	1315425	1315425	1315425	1315425	1315425
R2	0.000	0.000	0.000	0.000	0.000
R2 Adj.	0.000	0.000	0.000	0.000	0.000

Notes: This table reports OLS estimates of the coefficients from the regression:

$$\mathbb{1}_{\{win_i=1\}} R_i^* = \alpha + \beta R_i^* + \delta\, \text{OddsChange}_i + \gamma\, R_i^* \times \text{OddsChange}_i + \varepsilon_i.$$

The variable $\Delta R_{i,[-\tau,0]}/R_{i,-\tau} \equiv (R_i^* - R_{i,-\tau})/R_{i,-\tau}$ denotes the rate of change in odds over the final τ minutes before post time.

of residuals within each race. Clustering by race ensures valid inference under these forms of within-race dependence.

Estimation results incorporating these robustness checks are reported in Tables A.5. The results remain quantitatively and statistically robust, with the key coefficients of interest remaining highly significant and of similar magnitude to the baseline estimates.

Table A.5: Estimation Results with Race Fixed Effects: Winning Odds

	(1)	(2)	(3)	(4)	(5)
R_i^*	-0.00158	-0.00093	-0.00100	-0.00121	-0.00125
	(1.0122e-04)	(0.00020)	(0.00022)	(0.00021)	(0.00021)
$\frac{\Delta R_{i,[-5,0]}}{R_{i,-5}}$		-0.35756			
		(0.03946)			
$\frac{\Delta R_{i,[-10,0]}}{R_{i,-10}}$			-0.16842		
			(0.02844)		
$\frac{\Delta R_{i,[-15,0]}}{R_{i,-15}}$				-0.10831	
				(0.02453)	
$\frac{\Delta R_{i,[-20,0]}}{R_{i,-20}}$					-0.07297
					(0.02048)
$R_i^* \times \frac{\Delta R_{i,[-5,0]}}{R_{i,-5}}$		-0.00030			
		(0.00036)			
$R_i^* \times \frac{\Delta R_{i,[-10,0]}}{R_{i,-10}}$			-0.00023		
			(0.00029)		
$R_i^* \times \frac{\Delta R_{i,[-15,0]}}{R_{i,-15}}$				-0.00007	
				(0.00025)	
$R_i^* \times \frac{\Delta R_{i,[-20,0]}}{R_{i,-20}}$					-0.00008
					(0.00020)
Num.Obs.	894127	894127	894127	894127	894127
R2	0.054	0.054	0.054	0.054	0.054
R2 Adj.	-0.018	-0.018	-0.018	-0.018	-0.018

Notes: This table reports OLS estimates of the coefficients from the regression:

$$\mathbb{1}_{\{win_i=1\}} R_i^* = \alpha_{j(i)} + \beta R_i^* + \delta \, \text{OddsChange}_i + \gamma \, R_i^* \times \text{OddsChange}_i + \text{Race Fixed Effect} + \varepsilon_i.$$

The variable $\Delta R_{i,[-\tau,0]}/R_{i,-\tau} \equiv (R_i^* - R_{i,-\tau})/R_{i,-\tau}$ denotes the rate of change in odds over the final τ minutes before post time.