Proceedings of the 7th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/2025.AST29217

Stall Position Bias in British Horse Racing: A Comparative Analysis Across Distance and Course

Ziyu Lin[1], Xinyue Zhou[2], Yuxuan Wu[3], Xuan Shen[1*]

[1]*Schoolof Economics and Management, University of Electronic Science and Technology of China, Chengdu, China*
[2]*Faculty of Science and Engineering, University of Nottingham Ningbo China, Ningbo, China*
[3]*School of Mathematics and Statistics, University of Manchester, Manchester, Britain*
**Corresponding Author. Email: shenxuan0929@outlook.com*

Abstract. This study investigates the impact of starting draw position (inner vs. outer tracks) on race outcomes at three distinct UK racecourses—Southwell, Newmarket, and Pontefract—examining how race distance and course topography modulate this effect. Utilizing linear regression and Spearman Rank Correlation Meta-Analysis, this study quantified draw-related performance biases across an extensive dataset of historical races. Races were systematically categorized into short, middle, and long-distance groups by identifying inflection points in draw effect trends through graphical analysis. The findings demonstrate a statistically significant relationship between stall number and finishing position across all three courses, with the magnitude and direction of bias varying by distance and track layout. Key results reveal that inner draws consistently provide an advantage in short-distance races, likely due to reduced travel distance on tighter turns. Middle-distance races exhibit pronounced course-dependent biases, with outer draws particularly disadvantaged at Newmarket, possibly due to its wide, sweeping turns. In contrast, long-distance races show diminished draw effects overall, though outer draws surprisingly improve outcomes at Southwell, suggesting terrain-specific adaptations. Meta-analysis further confirmed significant heterogeneity in draw biases across courses, highlighting the interplay between distance-governed turn frequency and unique topographic features. These findings underscore that track position bias is not universal but shaped by complex interactions between race dynamics and course design. The study provides actionable insights for strategic draw selection and proposes adjustments to handicapping models to account for these biases, ultimately enhancing competitive fairness in horse racing.

Keywords: horse-racing, stall-number, finish-position, distance, courses, linear regression, Spearman Rank Correlation Meta-Analysis, Stall Position bias

1. Introduction

Horse racing is one of the most popular sports competitions, and it is simultaneously the world's largest betting marketplace by turnover: annual global handle now exceeds US $120 billion, with British racing alone accounting for roughly £4–5 billion each year [1]. The sport's wagering

DOI: 10.54254/2755-2721/2025.AST29217

ecosystem ranges from on-course bookmakers and high-street betting shops to algorithmic trading desks that update odds every few hundred milliseconds on Betfair and other exchanges. The vast majority of this money is staked on win and place markets that are exquisitely sensitive to tiny advantage that can swing expected value by double-digit percentages. Within this context, stall position has long been viewed by punters as a potential "hidden variable", which could be beneficial to secure the ideal start position and shorten the curve distance [2], but its real impact on finishing position remains contested.

Existing studies divide races into two classifications roughly - long and short, and basic on that, researchers found two different results in it. Early studies confirmed inner draws significantly benefit short-distance races less than 1400 meters [3]. Likewise, Carruthers notes that horses start from inner tracks conserve 3-5% energy early, dominating in sprints [4]. Furthermore, study shows stall saves 2-3% distance when meet sharp curves [5]. But outer paths enable tactical flexibility and this advantage diminishes beyond 1800 meters. Based on this finding, the interaction is confirmed that draw bias peaks at 5–7f, attenuates through 7f–1m, and becomes negligible beyond 1m2f, where stamina and class dominate [4].

After having a scanning read about these relative information, a common and general method that scholars most use is linear regression, which is similar to one of this study' s methodology. In their results, lane position is one factor among many influencing race outcomes in specific researching race courses. However, whether the stall effect persists across diverse race distances and tracks, or is modulated by these aspects instead, is yet underexplored through empirical analysis.

In order to fill this gap, the objective of this research is to prove the correlation between stall position and finishing rank, not only by linear regression but also Spearman rank correlation meta-analysis. This study will further explore it through evaluating how different race distance will influence draw efficacy and compare the race information of Southwell, Newmarket, and Pontefract to generate the draw-related advantages or disadvantages.

While stall position bias is often discussed anecdotally, robust empirical analysis across diverse courses and distances remains limited. This study addresses this gap by quantifying how draw position, distance, and course topography interact at three key UK venues. Our findings provide novel, actionable insights. Firstly, it offers jockeys, trainers, and organizers evidence-based strategies for draw selection tailored to specific race conditions; Secondly, it supplies handicappers with crucial data to adjust models for greater fairness by accounting for inherent stall advantages or disadvantages. Thirdly, it enhances prediction accuracy for bettors and analysts by clarifying when and where draw position significantly impacts results.

This paper is divided into five main sections to systematically present the research. The introduction presents the research question. The Data section details the comprehensive dataset utilized in this study, describing its source, scope, variables, and any preprocessing steps undertaken to ensure analysis readiness. The Methodology section outlines the robust analytical framework employed, specifically detailing the application of linear regression and Spearman Rank Correlation Meta-Analysis techniques to quantify the relationship between stall position and finishing rank. The Results section presents and interprets the key empirical findings, explaining the observed patterns of draw bias across the different distance categories (short, middle, long) and the three distinct racecourses (Southwell, Newmarket, Pontefract). The Conclusion section synthesizes the core insights and summarizes the study's contribution to empirical applications.

DOI: 10.54254/2755-2721/2025.AST29217

2. Data

2.1. Data acquisition and description

Data for this study were obtained from a racing website [6]. It includes the horse racing information of three representative courses, namely Southwell, Newmarket, and Pontefract, between 2003 and 2012.

The reason for choosing these three courses is that their distinct characteristics in size, distance, slope, and surface provide a diverse and representative sample, allowing for more generalizable conclusions. This selection strategy, considering such varied topography, addresses a limitation often present in past studies which focused on single or less diverse courses. Furthermore, the methodology employed, involving systematic distance categorization and robust statistical techniques (linear regression and Spearman Rank Correlation Meta-Analysis), offers a more rigorous and verifiable approach than many previous investigations.

Here is the brief introduction of these three distinct British racecourses:

Southwell is unique for its Tapeta fiber sand surface in the inner track, which is suitable for all-weather competition. This kind of surface allows horses to run at a consistent speed. Additionally, with rapid up and down hills as well as sudden curves, stall effect is reported to be significant. Figure 1 shows the stall's shape of Southwell.

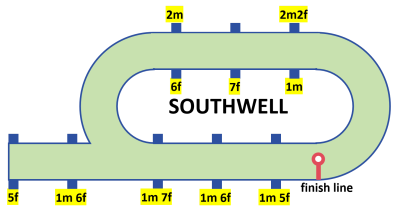


Figure 1. The stall's shape of Southwell

Newmarket is renowned for its wide and gently undulating turf land courses, including two independent tracks, the straight Rowley Mile and the bending July Course. This land requires horses to balance their speed and endurance to win the game. Figure 2 shows the stall's shape of Newmarket.

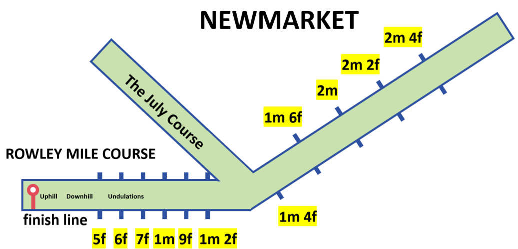


Figure 2. The stall's shape of Newmarket

Pontefract is the longest elliptical left-handed grass track in Europe. It has sharp turns and inward inclined runway, so the jokey need to adjust the position in advance. Figure 3 shows the stall's shape of Pontefract.

DOI: 10.54254/2755-2721/2025.AST29217

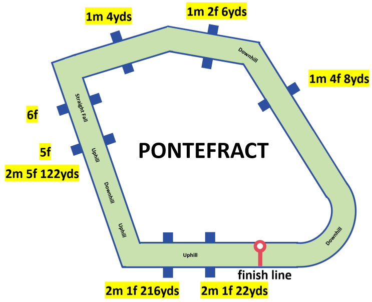


Figure 3. The stall's shape of Pontefract

2.2. Data preprocessing

To ensure the integrity and robustness of our statistical analysis, several data preprocessing steps were systematically applied. First, all entries with unknown track direction were removed, as this attribute is essential for evaluating lane-specific effects. Next, records of horses that did not finish the race are excluded, retaining only those explicitly marked as "finished". The distance-behind-winner metric was then standardized to enable consistent numerical comparison: entries marked as "/N", typically representing the winning horse, were replaced with zero, and any other non-numeric values were removed. These exclusions were consistent with prior filtering steps and helped establish a refined dataset for evaluating lane-based performance disparities.

To enhance the contextual understanding of race-level dynamics, a new variable, size, is introduced, representing the number of participating horses in each race. This value was computed by aggregating entries based on race id, allowing for systematic comparisons across races of varying field sizes. A cleaned subset of races from three distinct British racecourses were used to illustrate this framework. Entries with unknown track direction or non-finishing horses were excluded, and the distance-behind-winner field was cast to numeric after standardizing non-numeric values. To mitigate bias from atypically large fields, the study further excluded races with more than sixteen starters. To ensure comparability across races of different sizes, this study first assigned each horse a stall rank and a finish rank within its race using a first-tier rule. Then these ranks are converted into two continuous measures: fractional stall position and fractional finish position. This is done by dividing each rank by the total number of starters in that race. The fractional stall position and fractional finish position are each derived by dividing the respective stall rank and finish rank by the total number of participants in a race. As a result, both metrics are scaled to fall within the interval [0, 1]. This scaling enables direct comparison of positional effects across races with different numbers of participants and provides standardized inputs for all subsequent statistical tests.

Following data cleaning, exploratory analysis was conducted to uncover patterns in the relationship between track position (inner, middle and outer lanes). To gain initial insights into the relationships between stall positions and race outcomes across different tracks, this study conducted a structured exploratory data analysis. For each racecourse, the data are processed using the following steps. This study first grouped the dataset within each track by two key features. 'Distance yards' means the total distance of the race in yards. 'size' means the number of horses in the race. This ensured that the observations being compared shared similar race structures, reducing variability due to differing race formats. Each group therefore consisted of races with the same length and number of participants.

To better understand trends across varying race lengths, this study classified all distance yards values into three categories by the following steps. Firstly, plot draw effect metrics against race

DOI: 10.54254/2755-2721/2025.AST29217

distance for each course. Then, identify inflection points where the relationship between draw position and performance shifted significantly. And basic on this, the study separates these distances into 3 parts, (1) Short: Races below the first inflection point. (2) Middle: Races between inflection points. (3) Long: Races above the second inflection point. This stratification allows us to examine whether the influence of stall position on finishing position changes based on race length.

3. Methodology

3.1. Linear regression analysis

Linear regression is a fundamental statistical technique used to quantify the relationship between two continuous variables, in this context, the normalized stall number and the normalized finishing position. By fitting ordinary least squares regression models, the study aims to evaluate whether horses starting from certain stall positions tend to perform better (or worse) across various racing conditions.

OLS regression models were fitted for each distance-based subset within each course. The dependent variable was the fractional finishing position, and the independent variable was the fractional stall position. From the fitted models, the study extracted slope coefficients, standard errors, coefficient of determination (R^2), and p-values. To visually assess the nature and strength of these relationships, the study generated scatter plots with jittering applied to both axes, overlaid with regression trend lines. These plots help to illustrate whether the stall position effect varies systematically across different track layouts or distance groups. This approach allows for nuanced comparison of draw bias patterns across diverse race configurations, and helps identify whether certain stall positions confer systematic advantages under specific conditions.

3.2. Spearman rank correlation meta-analysis

To capture non-linear and rank-based associations across races, this study performed a Spearman correlation meta-analysis on the data. After reading in the long-format data set, this study grouped observations by race ID and calculated each race's Spearman correlation coefficient (ρ_i) between raw stall and finish ranks, retaining only races with at least three starters to ensure stable estimates.

Each (ρ_i) was converted to a Z-score via the formula $Z_i = \rho_i \,/\, \left(1/\sqrt{n_i - 1}\right)$, where n_i denotes the number of starters in race i . This fixed-standard-error approach standardizes correlation coefficients for aggregation. The overall effect size was then obtained as the mean of the Z_i values, and a two-sided p-value was computed under the assumption $Z_{\text{overall}} \sim N\left(0, 1/\sqrt{N}\right)$, where N is the count of included races.

Finally, this study produced a histogram of the individual Z_i scores with a standard normal density overlay. These diagnostic visuals confirm whether the assumptions underpinning the meta-analytic aggregation hold and provide insight into the variability of stall effects across individual races.

3.3. Methodological robustness considerations

The dual analytical framework employed, linear regression and Spearman rank-correlation meta-analysis, was deliberately designed to address inherent limitations in singular parametric approaches when modeling stall position effects. Whereas ordinary least squares regression quantifies the average linear relationship between fractional stall position and finishing rank, its validity hinges on

DOI: 10.54254/2755-2721/2025.AST29217

often-violated assumptions of linearity, homoscedasticity, and outlier insensitivity in complex racing data. The Spearman meta-analysis provides critical robustness by operating non-parametrically on rank-transformed variables, thereby immunizing estimates against three pervasive issues: extreme performance outliers (e.g., heavily favored horses overcoming disadvantaged draws), heteroscedasticity across variable field sizes, and non-linear threshold effects where advantages manifest only beyond specific stall positions. This methodological synergy proves particularly salient when interpreting discordant results: significant OLS coefficients with negligible Spearman correlations (e.g., Newmarket's short-distance races) suggest artifactually amplified linear trends from outliers or heteroscedasticity, while significant Spearman estimates with null OLS slopes (e.g., Pontefract's middle-distance subgroup) indicate monotonic but non-linear associations undetectable through parametric modeling. Consequently, the Spearman approach functions not merely as a supplementary technique but as a diagnostic validator that discerns structurally stable effects from statistical artifacts, ensuring conclusions reflect genuine biological or topological mechanisms rather than analytical idiosyncrasies.

4. Result

4.1. Race course: Southwell

4.1.1. Short distance races

For short distance races, shown in Table 1, linear regression analysis revealed a statistically significant positive association between the fraction of stall number and the fraction of finish position ($\beta = 0.1780$, $p < 0.001$), with an F-statistic of 160 on 1 and 4886 degrees of freedom. The positive slope indicates that horses starting from higher-numbered stalls tended to finish in worse positions. This finding is corroborated by the Spearman rank correlation meta-analysis, which produced an overall Z value of 0.504 ($p = 0$), suggesting a moderate positive association across races.

4.1.2. Middle distance races

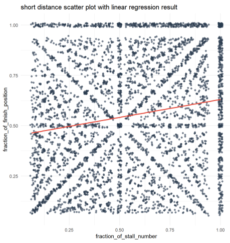
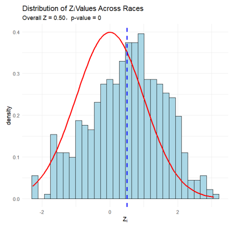
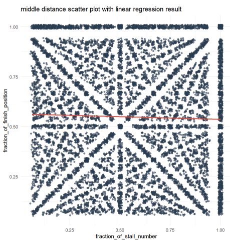
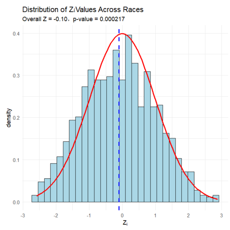
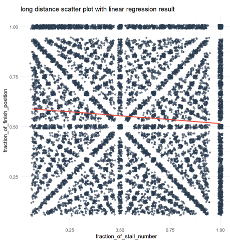
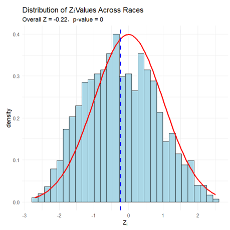
For middle distance races, shown in Table 1, the association between starting position and finishing position was much weaker. The linear regression slope was slightly negative ($\beta = -0.0268$, $p < 0.01$) with an F-statistic of 9.757 on 1 and 13,618 degrees of freedom. The small negative coefficient suggests that horses starting from higher-numbered stalls had a marginally better finish position, though the practical impact appears minimal. The Spearman analysis yielded an overall Z value of -0.101($p = 0.000217$), indicating a statistically significant but weak negative correlation.

4.1.3. Long distance races

For long distance races, shown in Table 1, a significant negative association was observed. The linear regression analysis showed a slope of -0.0782($p < 0.001$), with an F-statistic of 86.56 on 1 and 14,056 degrees of freedom. This suggests that in longer races, horses drawn in higher-numbered stalls tended to perform slightly better. The Spearman meta-analysis supported this finding with an overall Z value of -0.224($p = 0$), indicating a consistent negative correlation across races.

DOI: 10.54254/2755-2721/2025.AST29217

Table 1. The result of linear regression analysis and Spearman rank correlation meta-analysis of Southwell

	Linear Regression Analysis Result	Spearman Rank Correlation Meta-Analysis Result
short	fraction_of_finish_position =0.4525+0.1780*fraction_of_stall_number p-value < 0.001 F-statistic:160 on 1 and 4886 DF 	Overall Z = 0.504 p-value = 0 
middle	fraction_of_finish_position =0.5639-0.0268*fraction_of_stall_number p-value<0.01 F-statistic:9.757 on 1 and 13618 DF 	Overall Z = -0.101 p-value = 0.000217 
long	fraction_of_finish_position =0.5947-0.0782*fraction_of_stall_number p-value<0.001 F-statistic:86.56 on 1 and 14056 DF 	Overall Z = -0.224 p-value = 0 

4.2. Race course: Newmarket

4.2.1. Short distance races

For short distance races, shown in Table 2, linear regression analysis revealed a statistically significant positive association between the fraction of stall number and the fraction of finish position (β = 0.0019, p < 0.001), with an F-statistic of 86.56 on 1 and 14,056 degrees of freedom. The positive slope indicates that horses starting from higher-numbered stalls tended to finish in worse positions, although the effect size is relatively small.

In contrast, the Spearman rank correlation meta-analysis yielded an overall Z value of -0.004 (p = 0.875), suggesting no statistically significant monotonic relationship between stall and finish ranks across races. The discrepancy between the two methods may reflect the sensitivity of linear regression to small but consistent directional trends, whereas Spearman correlation captures only rank-based monotonicity.

DOI: 10.54254/2755-2721/2025.AST29217

4.2.2. Middle distance races

For middle distance races, shown in Table 2, the linear regression model produced a slightly negative slope ($\beta = -0.0073$, $p \approx 1$), with an F-statistic of 0.1469 on 1 and 2,789 degrees of freedom. This result indicates no meaningful linear association between stall position and finishing outcome. The negative coefficient suggests a marginal tendency for outer stalls to perform better, but the effect is statistically and practically negligible.

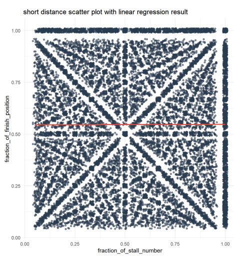
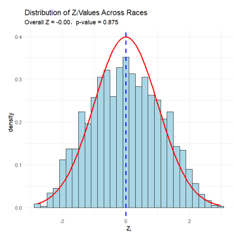
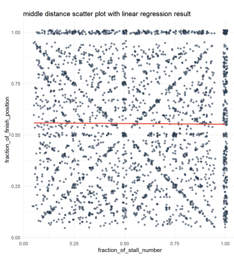
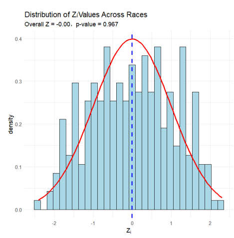
The Spearman meta-analysis supported this conclusion, yielding an overall Z value of -0.002 ($p = 0.967$), which confirms the absence of a significant rank-based correlation. Both methods consistently suggest that stall position has little to no impact on race outcomes in middle-distance events at Newmarket.

4.2.3. Long-distance races

For long distance races, shown in Table 2, the linear regression model revealed a slightly stronger positive association ($\beta = 0.0422$, $p \approx 1$), with an F-statistic of 0.7139 on 1 and 400 degrees of freedom. Although the slope is larger than in the other distance categories, the lack of statistical significance suggests that the observed trend may be due to random variation rather than a systematic effect.

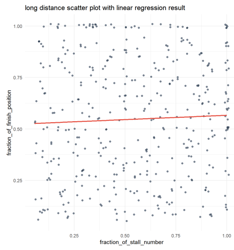
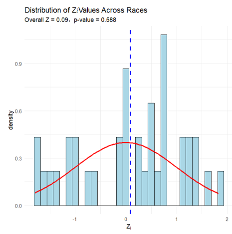
The Spearman rank correlation meta-analysis yielded an overall Z value of 0.089 ($p = 0.588$), again indicating no significant monotonic relationship. Taken together, these results suggest that stall position does not exert a consistent influence on race outcomes in long-distance races at Newmarket.

Table 2. The result of linear regression analysis and Spearman rank correlation meta-analysis of Newmarket

	Linear Regression Analysis Result	Spearman Rank Correlation Meta-Analysis Result
short	fraction_of_finish_position =0.5450+0.0019*fraction_of_stall_number p-value<0.001 F-statistic:86.56 on 1 and 14056 DF 	Overall Z = -0.004 p-value = 0.875 
middle	fraction_of_finish_position =0.5587-0.0073*fraction_of_stall_number p-value<1 F-statistic:0.1469 on 1 and 2789 DF 	Overall Z = -0.002 p-value = 0.967 

DOI: 10.54254/2755-2721/2025.AST29217

Table 2. (continued)

long	fraction_of_finish_position =0.5242+0.0422*fraction_of_stall_number p-value<1 F-statistic:0.7139 on 1 and 400 DF 	Overall Z = 0.089 p-value = 0.588 

4.3. Race course: Pontefract

4.3.1. Short distance races

For short distance races, shown in Table 3, linear regression analysis revealed a statistically significant positive association between the fraction of stall number and the fraction of finish position ($\beta = 0.1187$, $p < 0.001$), with an F-statistic of 90.03 on 1 and 6,300 degrees of freedom. The positive slope indicates that horses starting from higher-numbered stalls tended to finish in worse positions. This finding is corroborated by the Spearman rank correlation meta-analysis, which produced an overall Z value of 0.365 ($p = 0$), suggesting a moderate positive association across races.

4.3.2. Middle distance races

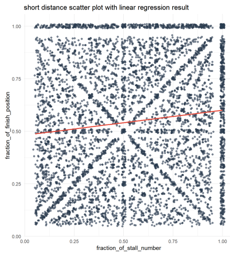
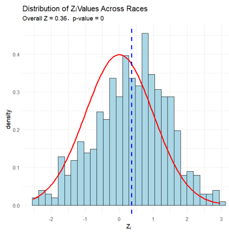
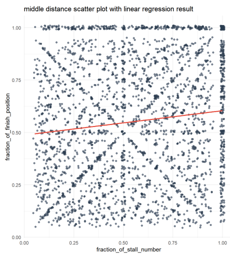
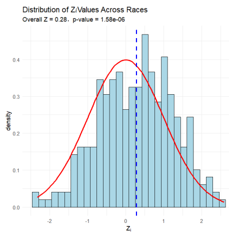
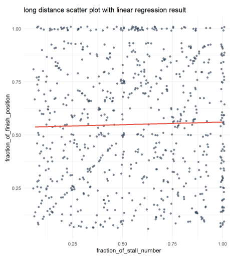
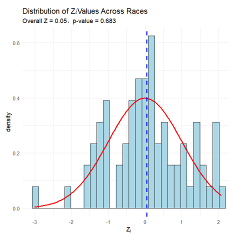
For middle distance races, shown in Table 3, the association between starting position and finishing position was weaker but still statistically significant. The linear regression slope was positive ($\beta = 0.1164$, $p < 0.001$), with an F-statistic of 40.9 on 1 and 2,976 degrees of freedom. The positive coefficient suggests that horses starting from higher-numbered stalls had a worse finish position. The Spearman analysis yielded an overall Z value of 0.282 ($p = 1.58e{-}06$), indicating a statistically significant positive correlation. Both methods consistently suggest that outer stalls are disadvantaged in middle-distance races at Pontefract.

4.3.3. Long distance races

For long distance races, shown in Table 3, the association between stall number and finish position was not statistically significant. The linear regression analysis showed a slope of 0.0241 ($p \approx 1$), with an F-statistic of 0.4285 on 1 and 736 degrees of freedom. This suggests that in longer races, the starting stall number had little impact on the finish position. The Spearman meta-analysis supported this finding with an overall Z value of 0.048 ($p = 0.683$), indicating no significant correlation across races.

DOI: 10.54254/2755-2721/2025.AST29217

Table 3. The Result of linear regression analysis and Spearman rank correlation meta-analysis of Pontefract

	Linear Regression Analysis Result	Spearman Rank Correlation Meta-Analysis Result
short	fraction_of_finish_position =0.4816+0.1187*fraction_of_stall_number p-value<0.001 F-statistic:90.03 on 1 and 6300 DF 	Overall Z = 0.365 p-value = 0 
middle	fraction_of_finish_position =0.4878+0.1164*fraction_of_stall_number p-value<0.001 F-statistic:40.9 on 1 and 2976 DF 	Overall Z = 0.282 p-value = 1.58e-06 
long	fraction_of_finish_position =0.5342+0.0241*fraction_of_stall_number p-value<1 F-statistic:0.4285 on 1 and 736 DF 	Overall Z = 0.048 p-value = 0.683 

5. Conclusion

This study provides a comprehensive empirical analysis of the relationship between stall number and finish position in horse racing across three distinct British racecourses (Southwell, Newmarket, Pontefract) and varying race distances. The key findings reveal that stall position significantly influences race outcomes, but its effect is highly context-dependent, modulated by track characteristics and race distance.

Track effects for each course are pronounced specifically. In Southwell, there exhibits the strongest stall bias, with clear distance-dependent patterns. For short races, inner stalls (high numbers) advantage horses. For long races, inner stalls confer a disadvantage. In Pontefract, there is significant advantage for inner stalls in short races, but minimal effect in longer distances. In

Newmarket, there is no consistent stall bias across any distance, affirming its reputation as a "fair" track where strategy and horse capability dominate.

The result shows that race distance will moderate draw bias. Short-distance races (≤1500 yards) consistently show strong negative impacts for outer stalls on curved or inclined tracks (Southwell, Pontefract). Beyond intermediate distances (>1800 yards), stall effects diminish or reverse (e.g., Southwell's long races). Both linear regression and Spearman meta-analysis yielded convergent results for significant effects (e.g., Southwell's short and long races, Pontefract's short races). Discrepancies in non-significant contexts (e.g., Pontefract's middle distance) highlight the value of multi-method approaches in detecting subtle or non-linear patterns.

Our research findings have practical application significance. For jockeys and trainers, strategic stall selection will influence the finishing rank on technical tracks. They can prioritize inner stalls for short races to minimize distance loss on curves. In long races at Southwell, outer stalls may offer tactical flexibility. At Newmarket, stall choice is less consequential, they just need to focus on horse form and pace strategy. For betting Markets and analysts, bookmakers can develop dynamic pricing models that automatically adjust odds for inner stall horses in short-distance races at Southwell and Pontefract, while maintaining standard odds for long-distance races where stall positions show minimal impact. Notably, Newmarket's straight course requires no stall-based adjustments at all, keeping its traditional betting approach unchanged. For bettors, the research suggests favoring inner-stall horses on curved tracks for short races, looking for proven speedsters on curved tracks for long races, and focusing solely on current form at straight tracks. Additionally, the study provides the foundational logic for innovative betting tools, including stall-distance hedging systems and track-specific advantage indicators. It also enables the creation of early warning systems to detect when odds deviate significantly from expected values based on stall position advantages. These methods are projected to enhance long-term returns for disciplined bettors. To ensure fair competition, racecourses should modify sharp turns or implement handicap systems to balance stall advantages. However, our research still has the following limitations. Exclusion of non-finishers and weather or track conditions may omit confounding variables. Our result's universality still needs to be tested by expanding to global tracks.

References

[1] Wise Guy Reports, 2024. Horse Racing Market Analysis and Forecast 2032. https: //www.wiseguyreports.com/reports/horse-racing-market.
[2] Lo, V. S., & Bacon-Shone, J. (2008). Probability and Statistical Models for Racing. Journal of Quantitative Analysis in Sports, 4(2). https: //doi.org/10.2202/1559-0410.1103.
[3] Stokes, T., Bagga, G., Kroetch, K., Kumagai, B., & Welsh, L. (2024). A generative approach to frame-level multi-competitor races. Journal of Quantitative Analysis in Sports, 20(4), 365–383. https: //doi.org/10.1515/jqas-2023-0091.
[4] Carruthers, 2025. R. Pace and Draw Analysis: The Hidden Factors Affecting Horse Racing Results. https: //lightspeedstats.com/horse-racing/pace-and-draw-analysis/.
[5] Thiruvenkadan, A. K., Kandasamy, N., & Panneerselvam, S. (2009). Inheritance of racing performance of Thoroughbred horses. Livestock Science, 121(2-3), 308–326. https: //doi.org/10.1016/j.livsci.2008.07.009.
[6] Betwise, 2025. Betwise Racing Analytics. https: //www.betwise.co.uk/.
[7] At The Races, 2025. Southwell Course Guide. https: //www.attheraces.com/course-guides/Southwell.
[8] At The Races, 2025. Newmarket Course Guide. https: //www.attheraces.com/course-guides/Newmarket.
[9] At The Races, 2025. Pontefract Course Guide. https: //www.attheraces.com/course-guides/Pontefract.
[10] R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.2025.