Examining the Effect of Home Field Advantage on Umpire Decision Making and Differential Favorability Between American League and National League Teams

By May Mishima

PSYC-100: Introduction to Psychology

Professor C. Sanderson

December 2025

**Introduction**

Fans often assume and expect that sports officials make near perfect objective decisions, especially in professional sports where outcomes can have major consequences. However, human judgement is susceptible to social pressure, environmental factors, and personal perception, even in settings centered around neutrality. Sporting events provide a unique context for studying these influences as officials must make rapid decisions in front of loud crowds and in pressure-filled situations. Home-field advantage is one well known pattern that may reflect more than just athletic performance. Understanding how situational pressures shape decision-making in sports can offer insight into broader questions about fairness, bias, and human behavior in everyday life.

Research across several sports shows that officiating decisions are often shaped by the social context surrounding the event rather than being entirely objective. Abbate et al. (2025) investigated whether video review technology reduced referee bias in professional soccer. In this archival study, they compared matches before and after the introduction of Video Assistant Referee (VAR) technology, analyzing officiating patterns such as fouls, and penalties awarded to home and away teams (Abbate et al., 2025). Their findings showed that even with a system designed to improve decision accuracy, referees continued to favor home teams, particularly in games with strong crowd reactions. The researchers concluded that social pressure from spectators can subtly shape how officials interpret confusing plays, suggesting that home-field advantage is partly driven by psychological influences on decision-making. This work demonstrates that bias can persist even among highly trained professionals and provides a foundation for examining similar patterns in other sports contexts.

Similar psychological influences have been observed in research on Major League Baseball umpiring. Another study analyzed large-scale pitch-tracking data to investigate how MLB umpires judge borderline pitches and whether their decisions show systematic bias (Flannagan et al., 2024). In their study, the actual location of each pitch was compared with the call made by the umpire, allowing the researchers to assess accuracy and detect consistent patterns in missed calls. Their findings revealed that umpire judgements were significantly influenced by contextual factors such as pitch count and game situation, and that these biases persisted even as overall accuracy improved in recent seasons. The authors concluded that MLB umpire decision making is shaped by psychological and situational pressures even if the umpire is trying to be fully objective. These results reinforce the idea that officiating bias can appear even in highly standardized systems, motivating further examination of how such biases may vary across different baseball contexts.

The present study extends prior research by examining whether home-field advantage influences umpire decision-making differently in the American League and the National League. Unlike previous work, this study uses a quantitative measure of “favorability,” which reflects the extent to which umpire calls benefited the home team in each game. Prior research shows that officials are more susceptible to bias when crowd and game pressure is high, and these contextual influences can subtly shift how ambiguous calls are interpreted. Comparing favorability across leagues is important because the AL and NL differ in crowd environments, ballparks, and umpire diversity, making league-level bias possible. Recent trends further motivate this comparison: National League teams have shown higher average attendance in the last five seasons and have won seven of the last ten World Series championships, suggesting stronger fan engagement and greater performance visibility. These patterns imply that NL home crowds may exert more pressure on umpires or create environments that subtly encourage favorable calls. To address this question, the study analyzed archival data from MLB games and compared overall patterns of home team favorability between the two leagues. This method was designed to test the hypothesis that home teams in the National League would receive more favorable calls than home teams in the American League.

**Methods**

***Materials***

The data for this study was obtained from Umpire Scorecards, which is a publicly available website that compiles pitch-by-pitch evaluations of MLB umpire performance. The dataset used in this study includes all 2400 regular season MLB games from the 2025 season, as listed on the website. For each game, Umpire Scorecards provides home and away teams and a “favor” metric, which represents the difference in run expectancy impact resulting from incorrect umpire calls, measured in runs. In other words, it measures the impact that the umpire’s calls had on the home team compared to the away team.

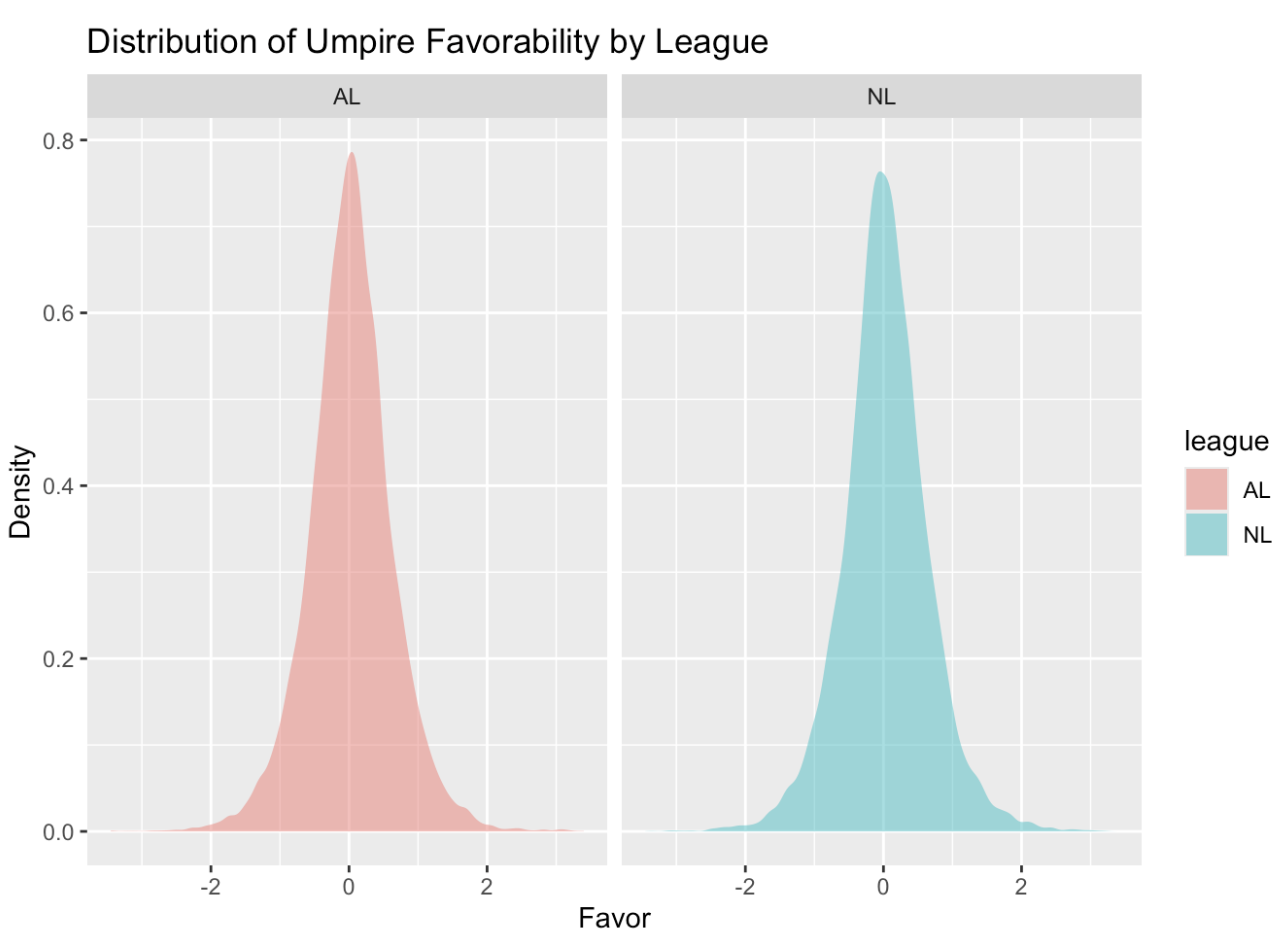
***Procedure***

The dataset was organized so that each game contained the home team’s favorability value and an indicator of whether the home team belonged to the American League or the National League. The primary variable of interest was favorability, defined as the run expectancy impact of incorrect umpire calls benefiting the home team relative to the away team. To evaluate whether the home teams in one league received more favorable calls than those in the other, favorability values were grouped by league and summarized using descriptive statistics. This included calculating the mean and variability within each league to assess overall differences in home-team advantage. Visualizations of the distributions were then created to examine whether patterns of favorability differed between the two leagues. Finally, statistical comparisons were conducted to determine whether the observed differences between AL and NL home teams were reliable.

The independent variable in this study was the league of the home team, which includes two levels, the American League and the National League. The dependent variable was the favorability that the home team experienced, based on the difference in run expectancy impact (home team - away team), measured in runs. A positive favorability indicated that the home team advantage was prevalent that game and a negative favorability indicated that the away team received more advantageous calls for that game. It is hypothesized that National League home teams received more favorability than American League home teams.

**Results**

It was hypothesized that home teams in the National League would receive more favorable umpire calls than home teams in the American League. The mean favorability for the American League home team was +0.026 runs, whereas the mean favorability for the National League home team favorability was +0.031 runs. Although both leagues showed favorability values centered near zero, the NL displayed a slightly higher average advantage for home teams. Thus, the hypothesis was supported: National League home teams received more favorable calls than American League home teams.



**Discussion**

The findings in this experiment indicate two major outcomes. First, home teams, regardless of league, received slightly more favorable calls than away teams, consistent with prior research on home-field advantage. Second, the American League and National League showed nearly identical favorability distributions, suggesting no meaningful league based differences in umpire bias. The similarity in means and distributional shape supports the conclusion that umpire decision-making does not systematically differ across leagues. Logically thinking, this is valid because umpires want to call games completely unbiased and go through intensive training to do so. However, home teams have crowds that may be louder for calls that would benefit their team, potentially influencing an umpire’s decision in the moment.

There were several limitations to this study that could have had a significant impact on the results. Firstly, the Umpire Scorecard dataset does not account for the psychological factors that may affect an umpire’s decision-making, such as crowd noise, game situation, weather conditions, stress levels, and personal biases towards specific players. These elements could make a meaningful difference in an umpire’s calls during a game. For example, louder crowds in a stressful game situation could cause umpires to increase favorability toward the crowd’s team. Since it was established in the introduction that NL teams typically attract larger crowds, omitting crowd defining variables likelyreduced the observed difference between the leagues. In other words, if crowd size truly amplifies home-field advantage, failing to include it almost certainly made the AL and NL appear more similar than they are under real game conditions. Together, these limitations suggest that true magnitude of league differences may be larger than reported here.

A second limitation of this study is the restricted time span of the dataset. Only the 2025 MLB season was analyzed, which limits the ability to detect long-term patterns in umpire favorability. Because league differences may fluctuate across seasons due to changes in crowd environments and weather/environmental factors, relying on a single year can systematically bias the results. In particular, if the 2025 season happened to be a year in which the American League and National League were unusually similar in average attendance and crowd intensity, then analyzing only this season would likely reduce true differences between leagues. Conversely, seasons with wider divergences in crowd size or more turnover in umpire rosters might show larger league differences than those observed here. The 2025 season does not appear to be atypical in any widely documented way, but without multi-year data it is impossible to determine whether it was a high- or low-variance year in terms of league-specific home-field environments. Therefore, using only one season likely made the present results less stable and potentially smaller in magnitude than what would be observed across multiple seasons.

Finally, the Umpire Scorecard dataset is based on observational archival data which limits the ability to draw causal conclusions. Since this was not a conducted experiment, it is impossible to attribute any patterns in favorability to the umpire's intended bias (conscious or unconscious) or differences in player behavior, catcher framing, pitch characteristics, and other random variation. Without the ability to isolate these factors, the present study cannot explain *why* the leagues appeared similar or what underlying mechanisms might shape favorability. This limitation highlights that archival analyses can reveal patterns in performance but cannot identify the specific psychological or environmental causes behind them.

One direction for future research would be to examine home team bias across the other major sports in North America, such as, NBA, NFL, NHL, and MLS. Each sport has different officiating structures, game paces, environmental factors and athlete-referee interactions, which may produce distinct patterns of favorability. Researchers can find similar datasets to the Umpire Scorecard or calculate relative favorability scores per sport, such as incorrect foul calls in the NBA, penalty yardage in the NFL, or missed penalty calls in the MLS and NHL. These metrics could be compared across leagues to determine whether officials in certain sports exhibit stronger home team advantage than others. One prediction is sports with closer fan proximity and louder crowds, like the NBA and NHL, may show larger favorability towards the home team than sports played in large outdoor stadiums such as the NFL or MLS. One logical explanation could be that closer physical proximity increases the immediacy and intensity of crowd reactions, heightening social pressure on officials in ways that may subtly influence judgement under uncertainty. Testing this prediction would be useful because it would reveal whether crowd-induced bias is a universal feature of officiating or a context-dependent phenomenon that varies by sport. Identifying these differences would help clarify which environmental and structural factors most strongly influence officials’ decision-making, extending the implications of the present study beyond baseball.

Furthermore, it would be valuable to examine “home field advantage” in a completely different setting, such as the legal system. Do judges favor and show more leniency toward defendants with whom they share some type of social commonality (e.g. hometown, education, demographic background)? Researchers could test this by archivally analyzing a random sample of cases and their sentences from one judge, identifying whether each judge and defendant share any measurable characteristic. The outcomes of these cases could then be compared across conditions where judge–defendant similarity is present versus absent. This process can be repeated with numerous judges to create a generalizable conclusion. The independent variables would include whether a similarity exists and the type of similarity, as well as the degree of the trial (e.g.misdemeanor, felony) and the dependent variable would be the severity of the sentence. If judges do indeed grant lesser punishments to defendants with shared backgrounds, this would suggest that in group bias extends beyond sports and influences high stakes real world judgements. The in group bias could be explained by the idea that humans naturally gravitate and empathize for those who have relatable experiences compared to those without any of those common characteristics. Both this present study and this potential future study explores whether subtle interpersonal similarities influence objective decision-making across very different contexts.

More broadly, the present study underscores the importance of continuing to investigate how social and environmental factors influence decision-making in professional sports. Even when league-level differences are minimal, the persistence of slight home-team advantages highlights how human judgment can be shaped by contextual pressures such as crowd reactions, situational intensity, and perceptual uncertainty. Research in this area is essential for understanding when and why officials may be vulnerable to bias, as well as for identifying the situational features that amplify or reduce these effects. Such work has meaningful real-world implications: it informs how organizations can design training, evaluation systems, and accountability structures that support more consistent and impartial decision-making. Continued research across seasons, leagues, and sports will help clarify the psychological mechanisms underlying officiating bias and contribute to building fairer, more transparent systems in sports and other high-stakes environments..

**References**

Abbate, C., Cross, J., & Uhrig, R. (2025). Video assistant referee and home field advantage: Implications for referee bias. *Southern Economic Journal*, *91*(3), 1176–1196. https://doi.org/10.1002/soej.12731

Flannagan, K. S., Mills, B. M., & Goldstone, R. L. (2024). The psychophysics of home plate umpire calls. *Scientific Reports, 14(1), 2735*. https://doi.org/10.1038/s41598-024-52402-y