The Effect of Missed Calls on Pitching Strategies

Brandon Yi, Jai Shah, Greg Morris

Introduction

In Major League Baseball, the accuracy of umpire strike zone calls has a large impact on the dynamics between pitchers, batters, and catchers. Now, tracking technology enables us to detect when strike zone errors are made by umpires, however, the strategic effects of these errors have yet to be studied. Specifically, this report investigates the extent to which incorrect umpire strike calls influence subsequent pitching strategy-focusing on how pitchers and catchers adjust pitch location after a missed call.

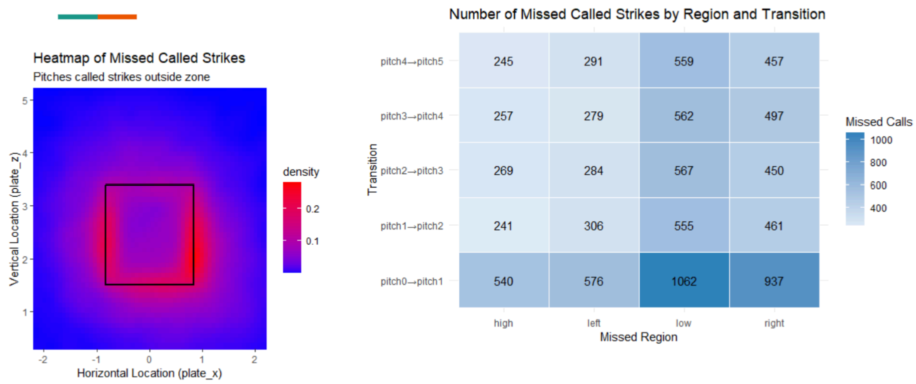
This question is especially applicable now as MLB moves toward implementing a challenge system for balls and strikes. Being able to use data analytics to increase understanding how missed calls affect pitch sequencing and location could provide teams with valuable insight, informing when to challenge a call and how to adapt in real-time. For instance, if a low strike is incorrectly called, pitchers may try to target that lower zone more aggressively, potentially forcing batters into less favorable swings and increasing the likelihood of ground balls in double play situations. Knowledge of these patterns could help batters and teams make more strategic challenge decisions, maximizing their chances of success in crucial moments.

Our analysis aims to quantify the relationship between missed umpire calls and subsequent pitching behavior, offering insights that could influence both in-game tactics and broader league policy. From our research, we were able to find that missed umpire calls have a significant impact on the pitching strategy of the pitcher, as well as the location of the subsequent pitches. In fact, we found that following a missed call, pitchers are more likely to place the ball in a similar location in subsequent pitches.

Data

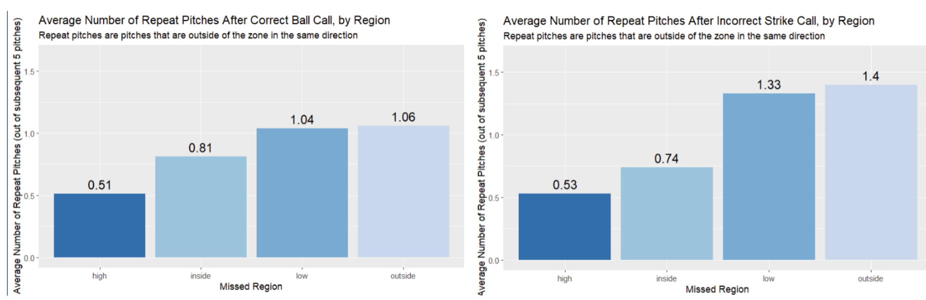
The data for this analysis was sourced from Baseball Savant and consists of detailed pitch-by-pitch records from Major League Baseball games. Each row in the dataset represents a sequence of pitches within an at-bat, capturing the location, outcome, and context for an initial pitch (pitch0) and up to five subsequent pitches (pitch1 to pitch5). For each pitch, variables include horizontal and vertical location (plate_x, plate_z), the umpire's call (description), the batter-specific strike zone boundaries (strike_zone_top, strike_zone_bottom), and contextual details such as game ID, pitcher ID, and handedness of both pitcher and batter. This structure allows us to precisely identify when an umpire makes an incorrect strike call-defined as a called strike outside the official strike zone-and to track how pitchers adjust their targeting in the following pitches.

EDA 1



Exploratory data analysis (EDA) highlights several key patterns relevant to our research question. As shown in EDA 1, a heatmap of missed called strikes reveals that most missed strikes occur just outside the formal strike zone, especially low and to the sides, with a transition matrix showing that the majority of these missed calls happen on the first pitch after the initial error and are most frequent in the low and right regions.

EDA 2



EDA 2 further demonstrates that after an incorrect strike call, pitchers are more likely to repeat pitches in the same out-of-zone region, particularly low and outside, with the average number of repeated pitches rising from around 1.0 after a correct ball call to as high as 1.4 after a missed strike. These findings suggest that missed calls not only cluster in specific areas but also have a measurable effect on subsequent pitch location, supporting the relevance and quality of the dataset for analyzing the influence of umpire error on pitching strategy

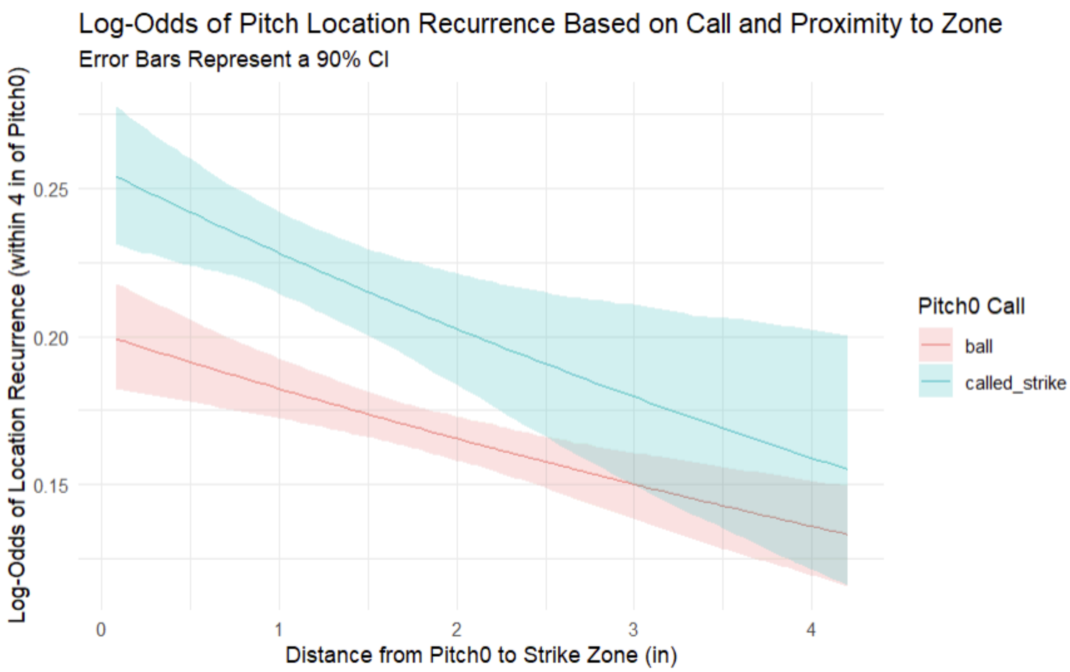
Methods

The first model, a standard logistic regression model, examines how pitch call type (ball or called strike) and proximity to the strike zone influence the log-odds of location recurrence in subsequent pitches. This model makes several key assumptions: (1) independence of observations, assuming each pitch-outcome pair is unrelated to others after conditioning on predictors; (2) a linear relationship between predictors and the log-odds of the outcome; (3) no severe multicollinearity among predictors; and (4) adequate sample size relative to the number of parameters. We specifically model the log-odds that a subsequent pitch appears within 4 inches of the initial pitch location as a function of the initial call and its distance from the strike zone edge as well as an interaction term between these two predictors. This approach is appropriate for our research question because it directly quantifies how umpire decisions affect pitch location choices, while the logit transformation accommodates the binary nature of our outcome. Uncertainty is quantified through standard errors derived from the information matrix, with 90% confidence intervals visualized as shaded bands around the regression lines.

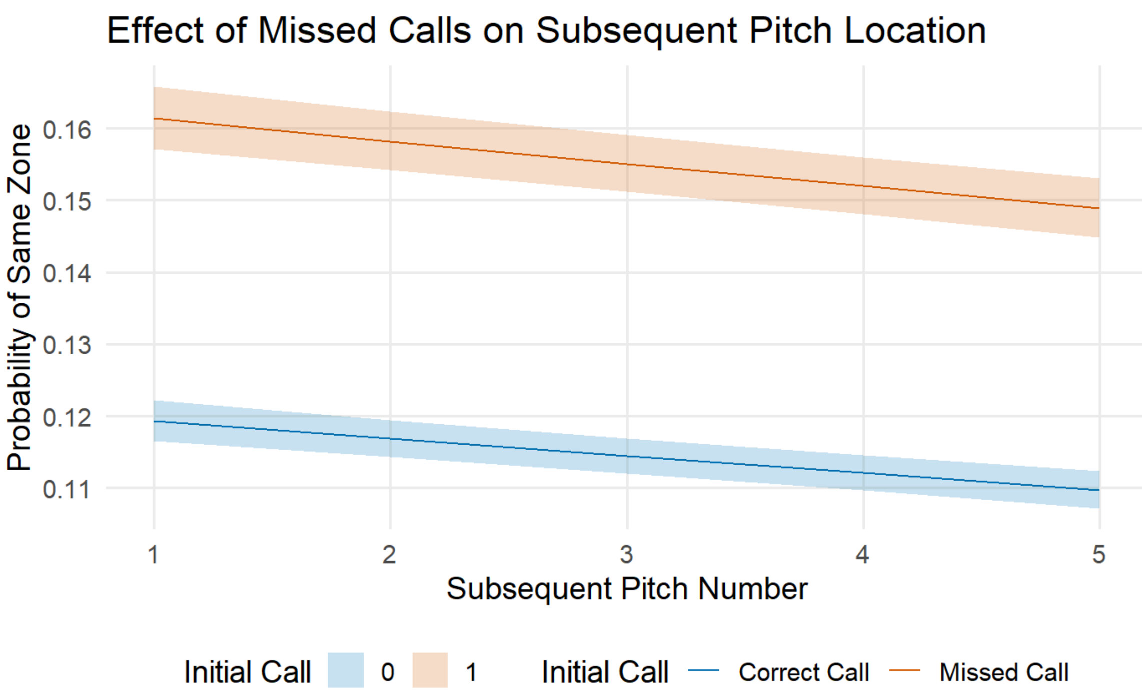
The mixed-effects logistic regression extends our analysis by accounting for the hierarchical structure of baseball data, where pitches are nested within pitchers, games, and at-bats. This model makes additional assumptions beyond standard logistic regression: (1) observations within groups (e.g., pitches thrown by the same pitcher) are correlated; (2) random effects follow a normal distribution with mean zero and estimable variance; (3) conditional on the random effects, observations are independent; and (4) both fixed and random effects have a linear relationship with the log-odds of the outcome. We model the probability that each of the next five pitches targets the same zone as the initial pitch, with a primary fixed effect for whether the initial call was missed or correct, and random intercepts for pitcher, game, and pitch sequence. This hierarchical approach is especially appropriate because it captures both population-average effects of missed calls and the natural variability in how individual pitchers adjust their strategies.

Both models enable different but complementary insights: the standard logistic regression reveals how the effect varies with pitch location relative to the strike zone, while the mixed-effects model captures how the effect persists across a sequence of pitches and varies across different contexts. For model evaluation, we compare AIC/BIC values to assess relative fit and use likelihood ratio tests to determine the significance of random effects. We also examine the predictive accuracy through confusion matrices and ROC curves. The mixed-effects approach provides a more robust uncertainty quantification by partitioning variance into within-cluster and between-cluster components, acknowledging that observations from the same pitcher or game are not truly independent. This comprehensive modeling strategy allows us to isolate the strategic impact of umpire errors while addressing the complex dependencies inherent in sequential pitch data.

Results



The logistic regression model, as displayed in the graph above, reveals that the umpire call for the first pitch in the sequence has a significant effect on the probability that a pitcher throws at least 1 pitch out of the subsequent 5 pitches in a similar location to the initial pitch. The model coefficient corresponding to the umpire call had a p-value of 0.00242 which is below our 0.05 significance level suggesting that the umpire call does indeed have a statistically significant effect on the pitcher's location strategy. Additionally, the distance that the initial pitch was from the strike zone does have a significant effect on the subsequent pitch locations with a p-value of 0.00165 and a coefficient of -1.40014. This suggests that as the initial pitch moves further away from the zone, pitchers are less likely to throw subsequent pitches in a similar location. This makes sense for called balls since pitchers will try to avoid missing the strike zone after they are punished for poor pitch placement, but why is this the case with called strike? It might seem logical that a pitcher would want to utilize a very wide strike zone if the umpire is incorrectly calling strikes further outside of the zone. While this may be the simple interpretation of the logistic regression output, the confidence interval reveals why this may be. The graph shows that for called strikes far outside of the zone there is a wide confidence interval which arises due to the fact that umpires rarely make incorrect calls on pitches that miss the strike zone by a significant margin, thus there is not enough data to properly model the relationship between distance and similar pitch location frequency for these types of pitches. This uncertainty produces an unintuitive result that can be ignored for this analysis.



The mixed effect model shows that pitch locations tend to be repeated in close succession rather than later in the pitch sequence. This is especially true for pitches that are incorrectly called strikes. While the effect of position within the sequence has a minimal effect it is still statistically significant with a p-value for the corresponding coefficient of 0.032. The result is unsurprising since our pitch sequence data is built in a way that spans at bats for the same pitcher within the same game, meaning within one sequence of 6 pitches (initial pitch plus subsequent 5 pitches) there may be multiple at bats represented. Considering that pitchers and catchers often have scouting reports for each batter, the pitch location strategy may vary significantly within our 6 pitch sequence. This variance is reduced for pitches that are closer in sequence since they are more likely to be thrown to the same batter, whereas a pitch 4 pitches after the initial pitch is more likely to be thrown to a different batter. When analyzing the random effects for pitchers we noticed very minimal differences between individuals suggesting that pitchers utilize missed calls at a similar rate across the league. This could be the case for a variety of reasons including the high difficulty in identifying when pitches are incorrectly called strikes for players in real time and the shared avoidance of throwing multiple pitches in the same location. The graph also shows that the confidence interval for incorrectly called strikes is wider than that of correctly called balls which is a result of a far fewer number of incorrect calls in our dataset.

Discussion

Overall we learned that umpires have an influence on the game of baseball and how pitchers pitch. From our initial EDA we learned that a lot of the missed called strikes are within 3-4 inches of the strike zone (measured from the middle of the baseball) and the most common areas that are repeatedly targeted following missed calls are low and outside. Then from our model we concluded that the closer an incorrect strike is to the strike zone the more likely a pitcher is to go back to the same spot within the next 5 pitches typically within an at bat.

During this project we faced limitations on the fact that we only used one month of data to try and conclude how pitchers pitch. Another limitation is that some of the data in the data set was null meaning the outcome of the pitch was not recorded so this went from an already small dataset to an even smaller one. The next steps for this research question include factoring in the umpire as some umpires are more lenient with certain areas of the strike zone. Another step would be to factor in the count of the at bat because you see pitchers throw different pitches and aim for different parts of the strizone based off of the count. For example, if the count is 3-2 or 2-2 and the umpire just called a strike low and outside you would imagine a pitcher would go right back to the same pitch. As opposed to if the count is 0-2 the pitcher would more likely try throwing further off the plate to try and bait the batter to swing.

In summary, our analysis highlights the nuanced impact that umpire decisions have on pitcher behavior, particularly regarding pitch location following missed strike calls. While our findings suggest pitchers are more likely to throw in similar locations near the strike zone after a borderline call, the scope of our conclusions is limited by the size and completeness of our dataset. Despite these constraints, our results provide a foundation for further exploration into how both umpire tendencies and situational factors, such as pitch count, shape pitching strategies. Moving forward, expanding the dataset to include a broader time frame and incorporating additional variables like individual umpire profiles and count context will allow for a more comprehensive understanding of these dynamics. Ultimately, this work lays the groundwork for more robust models that can better capture the complex interplay between umpires, pitchers, and the evolving strategies within the game of baseball.