# Learning Curves: Examining Post-Trade-Deadline Pitcher and Catcher Relationships

Berkeley N. Ho    Leonardo Robles-Lara    Christian D. Franco

Hayden A. Dunn

March 15, 2025

## 1 Research Question and Context

### 1.1 Motivation

In professional baseball, communication between pitchers and catchers is fundamental to team success. These duos are the heart of a team's defensive strategy, working in unison to prevent opposing hitters from converting plate appearances into runs. Generally, pitchers are responsible for metrics related to ball-control, including their velocity, spin rate, and movement. Notably, catchers have the crucial role of framing pitches, which can significantly influence the outcome of an at-bat. In optimal situations, catchers who frame well can reduce the number of pitches a pitcher has to throw, while less skilled catchers can have the opposite effect. Additionally, catchers "call games," or dictate which pitch type to throw in a given situation. Although scouting reports and coaches provide a game plan on how to face batters, it is up to both the pitcher and catcher to execute any in-game adjustments.

Before the regular season starts for Major League Baseball (MLB), spring training games provide a competitive-but-low-stakes setting for pitchers and catchers to acclimate to each another. However, the League's annual Trade Deadline provides a unique opportunity for analysis. It marks the final date when teams can trade players, typically occurring with about one-third of the season's games remaining. Players who are traded are then expected to immediately relocate and play for a new team, who have often brokered deals to supplement or fill roles necessary when competing for a playoff spot. Thus, the post-trade period acts as an observational study in which we can analyze how pitcher-catcher relationships develop in a more high-pressure context.

### 1.2 Research Question

Presently, there exists sociological literature on team chemistry in a variety of contexts, including sporting and non-sporting fields. Within sport settings, research has been conducted to model relationships via networks for both the men's National Basketball Association (NBA) and MLB. In the MLB, Wins Above Replacement (WAR) has been used to model the best arrangement of players. While the use of empirical analysis in baseball has grown exponentially since the sport's inception, there still remains a gap in the literature in formally quantifying or attempting to measure chemistry between individuals directly, including pitchers and catchers.

Thus, the aim of our project is to determine a method to quantify how "chemistry" develops via pitcher-catcher performance after trade deadlines.

### 1.3 Objective

As part of our goal in quantifying chemistry between pitchers and catchers, we have several objectives:

- Can we determine what variables are most correlated to performance on a short and long term performance basis?

- Can we produce a measurable statistic on how acquiring a player would impact a team's immediate performance vs. games down the stretch?
- Can we provide a timeline for when peak performance or chemistry is first attained, or produce a curve of how chemistry develops across time?

# 2 Data Sources

To investigate these objectives, we will be using data aggregated from several sources:

Data was collected from recent pitchers or catchers who were traded after the start of the season as examples of relationships that do not yet have "chemistry." This will operate under the assumption that those pitchers and catchers have not played together before in any professional context.

Ideally, data post pitch-clock-implementation (start of 2023 season) would provide enough of a sample size for pitchers and catchers. However, to increase our sample size, data was collected for the 2021-2024 seasons. It would be important to also consider other major rule changes or conditions, including how the composition of baseballs in recent seasons has led to discussions on whether balls are "juiced", leading to changes in batting performance compared to previously. As a result, pitcher and catcher performance should be sampled with these changes in mind.

To create our player dataset, we used MLB.com for an accurate record of all players involved in trades that were triggered one month before the trade deadline from 2021 to 2024. We web-scraped transactions from these seasons using the Python package `Beautiful Soup`. From the transaction webpages, we extracted the trade date, player name, handedness, and teams involved in the trade.

Baseball Savant, Baseball Reference and MLB.com all provide web based sources from which to subset and pull data for these players. In order to access this data, we used the R package `baseballr`. The `statcast_search_pitchers()` and `pitcher_game_logs_fg()` functions were primarily used to sort through pitcher data by their unique identification numbers in order to access their pitch-level data. This approach enables us to track specific games and players involved, including catcher IDs, with game logs sourced from FanGraphs. This framework sets the foundation for accessing the data necessary to answer our research questions.

Trackman data from each pitch, along with various other data points, will enable us to build a model that incorporates multiple statistics to quantify pitcher and catcher relationships and their evolution over the course of a season.

Performance will differ based on what type of role a pitcher holds. Accordingly, starters and relievers/closers will be analyzed separately. Relievers and closers, with fewer total innings pitched but more frequent appearances, typically throw multiple times per week, in contrast to starters. Starting pitchers were typically characterized as starting games and having more innings per outing on average. As a result, starters will be designated if they average more than 4 innings per outing in the first 7 innings of the game.

By clustering pitchers, we will further analyze how each performs relative to their peers, incorporating important metrics like 90th percentile fastball velocity and average innings per outing.

In order to distinguish the impact of pitcher and catcher performance on a team from that particular pitcher-catcher duo's chemistry, we have engineered features to quantify measures of familiarity. These include the number of games that pitcher and catcher have played in together, the number of days since being traded, and batters faced together.

# 3 Exploratory Data Analysis

Table 1 presents a comparison of pitcher performance before and after a trade, segmented by role (relief pitcher vs. starting pitcher).

For relief pitchers, we observe a slight decrease in innings pitched (IP) after the trade, from 1.22 pre-trade to 1.17 post-trade. Earned runs (ER) show a decrease as well, from 0.83 pre-trade to 0.66 post-trade, indicating an improvement in performance post-trade. The K/9 rate remains relatively stable (9.23 pre-trade vs. 9.32 post-trade), suggesting that the strikeout ability of relief pitchers is not notably affected by the change in team or catcher. There is a slight improvement in BB/9, from 4.50 pre-trade to 4.30 post-trade, and a decrease in WHIP (from 1.72 to 1.53), indicating better

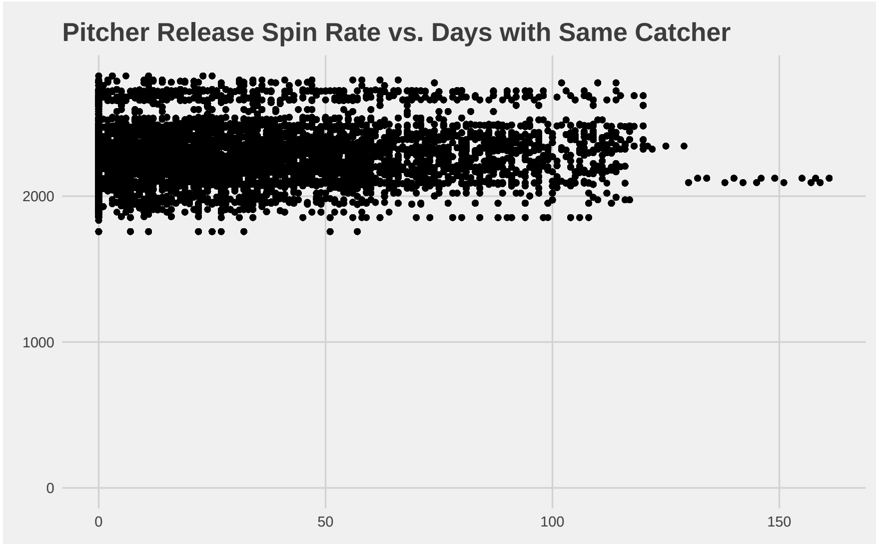


Figure 1: Initial scatterplot of spin rate as number of days with a catcher increases

control and fewer base runners allowed. The FIP remains relatively similar, showing that the pitcher's performance, independent of fielding, doesn't change dramatically post-trade.

For starting pitchers, there is a noticeable difference in innings pitched (IP), where the post-trade figure (4.99) is slightly lower than pre-trade (5.21). Despite this, earned runs (ER) are slightly higher post-trade (2.72) compared to pre-trade (2.48). However, K/9 decreases from 8.17 pre-trade to 7.63 post-trade, indicating that the strikeout rate for starters declines after the trade. There is a slight increase in BB/9, from 3.63 to 3.90, and the WHIP remains fairly consistent (1.49 to 1.53). BABIP and FIP also exhibit minimal changes, suggesting that the fielding-independent performance of starters remains stable, even if their overall effectiveness (as measured by earned runs) has decreased slightly post-trade.

This initial summary highlights the potential for trade-related changes in performance, especially regarding pitch control and strikeout ability. The data suggests some improvement in relief pitchers' control and earned run average post-trade, while starting pitchers show some decreases in strikeouts and a slight uptick in earned runs. These observations will be useful as we continue our analysis and attempt to identify more significant patterns or correlations with pitcher-catcher chemistry. It is also important to note that any differences in performance could be random, or attributable to other factors such as fatigue.

| Trade Status | IP | ER | WP | K/9 | BB/9 | WHIP | BABIP | FIP | Pitches Thrown | Role |
|---|---|---|---|---|---|---|---|---|---|---|
| Post-Trade | 1.17 | 0.66 | 0.06 | 9.32 | 4.30 | 1.53 | 0.26 | 4.71 | 22.11 | Relief |
| Pre-Trade | 1.22 | 0.83 | 0.06 | 9.23 | 4.50 | 1.72 | 0.27 | 5.32 | 23.48 | Relief |
| Post-Trade | 4.99 | 2.72 | 0.11 | 7.63 | 3.90 | 1.53 | 0.28 | 5.32 | 84.98 | Starting |
| Pre-Trade | 5.21 | 2.48 | 0.12 | 8.17 | 3.63 | 1.49 | 0.30 | 4.53 | 87.77 | Starting |

Table 1: Pitcher Performance Before and After Trade for Relief and Starting Roles

Both Figures 1 and 2 appear to have noise due to the number of observations. In future analysis, it would be better to look at the groups separated between relief and starting pitchers. Additionally, it will be important to have a better framework for summarizing individual performance and how that might impact overall trends with pitchers appearing multiple times in the same time period. For future analysis, limiting observations to having a qualified amount of pitches above a certain amount might produce more clear results.

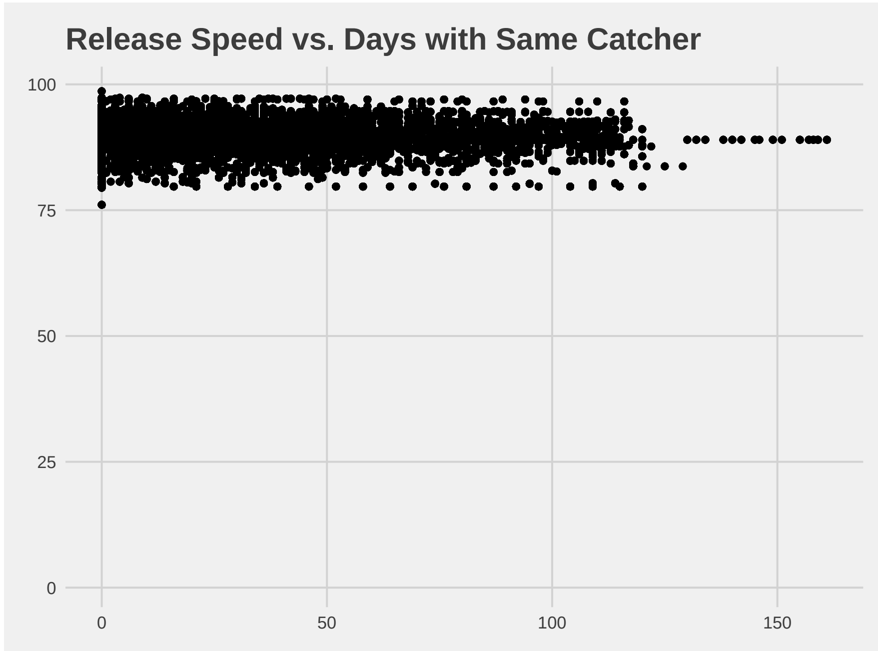


Figure 2: Initial scatterplot of release speed as number of days with a catcher increases

# 4 Methodology

To quantify how ”chemistry” develops between pitchers and catchers after the trade deadline, we will employ a combination of linear regression and machine learning techniques, including XGBoost and ensemble methods. These methods are well-suited to model relationships in our data and provide insights into the performance dynamics between pitcher-catcher duos.

Linear regression will be employed to examine the relationship between pitcher-catcher performance metrics and their influence on overall team outcomes. We will focus on the changes in key performance variables before and after the trade and explore how the duration of time spent together impacts these changes. Specifically, the models will quantify how performance metrics such as velocity, strikeout rate, and pitch usage correlate with the time a pitcher spends with a new catcher post-trade.

This method will allow us to quantify the influence of individual variables, such as pitch velocity, framing runs, and strikeout rates, on outcomes like ERA and team win-loss records after trades. By focusing on both short-term and long-term performance, linear regression will help us identify which metrics are most predictive of success following the trade deadline. The model will be set up to predict performance changes over time and test for statistical significance, providing clarity on which factors contribute most to the development of ”chemistry.”

We will also use machine learning techniques, such as XGBoost and ensemble methods. XGBoost is more adept at handling complex, non-linear relationships in large datasets. This will help us identify interactions between pitcher and catcher metrics that linear models might overlook. Ensemble methods will be utilized to combine multiple models, improving prediction accuracy by reducing overfitting. These models will be trained on various performance metrics to predict outcomes such as ERA improvement, strikeout rate increase, and other relevant variables over time. Additionally, we will use feature importance scores from the machine learning models to identify which performance metrics contribute most to the ”chemistry” between pitchers and catchers.

Additionally, we will use model validation techniques, such as cross-validation, to assess the generalizability of our models. This ensures that our predictions on pitcher-catcher performance are reliable and not overly specific to the training data.

## 5 Next Steps

With data collection now complete, our next steps will focus on refining our methods and implementing the necessary models. Currently, we want to review or refine how to summarize pitching performance, deciding whether to use average performance across games or quantiles. This will be necessary for accurately reflecting pitcher consistency and variability in performance.

After making any necessary changes, we we will implement linear regression models to explore how various performance metrics (such as velocity, strikeout rate, and pitch usage) relate to the time spent with a new catcher. We also anticipate comparing the results of linear regression to more complex models, like XGBoost, to assess whether a machine learning approach yields more accurate predictions and insights into pitcher-catcher dynamics.

To evaluate the performance of these models, we will split our data into training and testing sets. The training set will be used to fit the models, while the testing set will allow us to assess model generalizability. Key metrics such as mean squared error (MSE), R-squared, and cross-validation will be used to compare the models' accuracy and ensure they are not overfitting.

Finally, to enhance the interpretability of our models, we will focus on creating visualizations such as SHAP plots. These will help us understand the impact of individual features on the models' predictions, providing more transparency and insight into which factors contribute most to pitcher-catcher performance.