

University of San Diego

Digital USD

---

Undergraduate Honors Theses

Theses and Dissertations

---

Spring 5-2018

## Location, Location, Location: Analyzing the Impact of Pitch Command on At-Bat Outcomes in Major League Baseball

Noah Hilton  
*University of San Diego*

Follow this and additional works at: [https://digital.sandiego.edu/honors\\_theses](https://digital.sandiego.edu/honors_theses)



Part of the [Sports Management Commons](#)

---

### Digital USD Citation

Hilton, Noah, "Location, Location, Location: Analyzing the Impact of Pitch Command on At-Bat Outcomes in Major League Baseball" (2018). *Undergraduate Honors Theses*. 94.  
[https://digital.sandiego.edu/honors\\_theses/94](https://digital.sandiego.edu/honors_theses/94)

This Undergraduate Honors Thesis is brought to you for free and open access by the Theses and Dissertations at Digital USD. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of Digital USD. For more information, please contact [digital@sandiego.edu](mailto:digital@sandiego.edu).

Location, Location, Location:  
Analyzing the Impact of Pitch Command  
On At-Bat Outcomes in Major League Baseball

---

A Thesis  
Presented to  
The Faculty and the Honors Program  
Of the University of San Diego

---

By  
Noah Hilton  
Business Economics  
2018

## Abstract

In baseball, runs are considered the currency of the game. As a result, the foremost value of pitchers to their respective teams comes in the form of run prevention and keeping the opposing offense off the scoreboard as much as possible. While Major League Baseball has taken steps to better evaluate the run prevention skills of pitchers independent of their defensive teammates, new radar technology has made possible the research of specific skills related to a pitcher's repertoire of offerings. In this article, the author first examines the relationship between various pitch characteristics — velocity, movement, and location — and the primary offensive-suppression objective of Major League Baseball pitchers. Evidence from the research suggests that characteristics related to velocity and control provide the greatest independent impacts on run prevention at the major league level. Building off of these results, the author then introduces a new statistic — the Indexed Command Diagnostic (ICD) — that is meant to empirically evaluate the pitchers who locate their pitches the best. Finally, the author examines the relationship between this new statistic and pitcher salaries in Major League Baseball to determine the efficiency of front office spending on big league arms, finding that baseball executives significantly undervalue meaningful pitch characteristics such as location relative to their respective on-field impacts. Collectively, the analyses may be used to determine that the implementation of the Indexed Command Diagnostic will aid not only run prevention efforts on the field by incentivizing productive pitcher behavior, but also assist in optimizing the financial commitments made by MLB organizations to big league pitchers off the field as well.

## Introduction

It is the classic mantra of real estate brokers across America: “Location, location, location.” At its core, the statement refers to a simple prioritization of the homebuyer’s wish list. One may find the perfect structure, with every imaginable amenity inside, but if the structure is poorly located, it becomes far more difficult to sell. The “where” is what truly matters to the consumer.

Similarly, location is an essential aspect of being a pitcher in baseball. From every Little Leaguer’s first game at the local sports park, it is drilled into the heads of aspiring arms by coaches, catchers, and well-intentioned parents. The logic seems sound: one can only get so far with velocity or movement, and a fastball down the heart of the plate is quickly found to be easier to hit than one that paints the outside corner. As a result, every young pitcher prioritizes three things: “location, location, location.”

That is, until that pitcher attempts to make the jump to the professional ranks. Suddenly, it seems the tools that are easiest to see - movement and especially velocity - are also rewarded most heavily. All 16 pitchers selected by teams in the first round of Major League Baseball’s 2017 First-Year Player Draft threw upwards of 90 miles per hour (mph), with an elite handful topping out at over 100 mph. The emphasis on radar gun readings doesn’t slow once players get into professional systems either, as evidenced by the fact that four of MLB’s five largest average-annual-value contracts signed over the last ten years went to power pitchers with fastballs in the mid-to-upper nineties — David Price, Clayton Kershaw, Max Scherzer, and Roger Clemens.

This system of incentives, along with advances in physical training and biomechanical analysis of pitching technique, has resulted in more MLB pitchers throwing with higher velocities than ever before. In 2008, 18 MLB pitchers threw fastballs in excess of 100 miles per hour, and these 18 pitchers combined for 196 fastballs of that velocity or higher. Fast forward ten seasons to 2017, however, and the league’s infatuation with high radar gun readings is clear: 40 pitchers combining for 1,017 total pitches of 100 miles per hour or faster (Gonzalez, 2018).

The unfortunate consequences of this systemically-encouraged need for speed are a rash of arm injuries — 36 torn ulnar collateral ligaments in Major League Baseball

since 2016, each requiring almost a full year of recovery — and an epidemic of less-than-optimal spending by MLB teams on pitchers who have the pure stuff to make scouts salivate, but who lack the polish to actually maximize the benefit of such overpowering repertoires, all while undervaluing and largely ignoring the potentially-equal contributions of arms that do not light up the radar gun with every pitch. It is a market inefficiency that has yet to be exploited.

This project hopes to do just that by beginning with the answer to one simple question: how important is location in the first place?

## **Literature Review**

Discussions of the proper methodology behind valuing baseball players date back to the early days of baseball researcher Bill James, who first coined the phrase “sabermetrics” to describe a discipline that uses statistics to analyze the machinations of America’s National Pastime. James’ *Baseball Abstract*, published annually since 1977, opened the door for others to further explore the weaknesses of traditional, commonplace statistics. One of these statistics is earned run average (ERA), which should theoretically be influenced to some extent by the skill of the defensive players asked to field grounders and fly balls behind the pitcher on the mound. Noting these flaws in one of the foremost ways those around the league evaluate a pitcher’s ability to prevent runs, researchers have since built off James’ studies to better evaluate pitcher skill and, consequently, to focus in on the performance measurements that most deserve to be factored into the financial compensation of major league arms.

Scully (first explored in 1974 and updated in 1989) was the first to take the research reins, using pitchers’ strikeout-to-walk ratios as a proxy for pitching skill and justifying his choice with the fact that the statistic performs well as a predictor of the winning percentages of teams. Zimbalist (1992) attempted to build on Scully’s model by utilizing ERA, which he found improved on strikeout-to-walk ratio as a predictive model of team performance, to better estimate the marginal revenue products (MRP) of MLB pitchers. Kahn (1993) and Krautmann (1999) also disputed Scully’s method, using their results to show that player compensation depends as much on player service time (and the

arbitration-eligible and free agent designations that accompany expanding service time) as it does on on-field statistical performance. These studies used ERA, along with a handful of other performance-based metrics, to estimate pitcher contributions toward winning and, consequently, organizational revenue.

However, studies since have disagreed with the findings of Kahn and Krautmann. Research performed by Voros McCracken (2001) expanded on Scully's methodology, hypothesizing that outcomes capable of pitcher production without the help of defensive players — walks, strikeouts, hit batters, and home runs allowed — provide an opportunity to evaluate pitching talent and skill independent of defensive influence. As a result, McCracken (2001) referred to these statistics as defense-independent pitching statistics (DIPS). He found that knowledge of how a pitcher performs without the assistance of his defenders provided nearly all of the information necessary to evaluate that pitcher's contributions to run prevention and team wins. While recognized in baseball's public analytics community, McCracken's research lacked verification by the academic and economic communities, and therefore required more formal scrutiny using appropriate econometric techniques.

JC Bradbury (2007) provided this scrutiny, analyzing the impacts of various pitching statistics on measures of run prevention and ultimately developing results that supported McCracken's research. Bradbury's findings showed that DIPS have higher correlations from season to season (suggesting the involvement of skill rather than luck or other external influences in their accumulation) than statistics involving balls in play. The results also lent evidence to the fact that DIPS serve as better indicators of run prevention ability than more traditional measures of success on the mound. Furthermore, Bradbury determined that Major League Baseball front offices had grown to recognize this in the years leading up to the publishing of Bradbury's research, suggesting that franchises were not only aware of these statistical reviews but were implementing them directly into their decisions on pitchers' salaries well before the release of McCracken's research.

Those findings corresponded with research done by Hakes and Sauer (2006) in the years following the release of Michael Lewis' *Moneyball* (2003), which focused on the Oakland Athletics' attempts to build a cost-effective roster through the expanded use

of more effective offensive statistics. Hakes and Sauer found that the baseball labor market for hitters had corrected its valuations of certain offensive statistics, including on-base percentage and slugging percentage, several seasons before the general public became aware of these erroneous valuations through *Moneyball*.

In the years since the studies of Bradbury and Hakes and Sauer, Major League Baseball's implementation of statistical analysis in the player evaluation process has accelerated. Additionally, the league has added new methods of producing and aggregating on-field data through the development of radar-based systems such as Trackman, PITCHf/x, and Statcast. On the pitching side of the equation, these new tracking systems have introduced the baseball community to such process-based metrics as velocity, spin rate, pitch location, and horizontal and vertical break on pitches.

Research into the impact of these pitch characteristics on the effectiveness of specific pitch types, including change-ups (Pavlidis, 2013), curveballs (FOX Sports, 2015), and sliders (Sarris, 2015), found that regardless of pitch type, the most meaningful pitch characteristics related to higher swing-and-miss rates and higher groundball frequencies were similar across the board (these being vertical movement and change in velocity relative to a pitcher's fastball). Further studies by Pemstein (2015) found that location and spin rate (how quickly the ball rotates on its path to the plate) were also useful in explaining the types of batted balls that pitchers are able to induce. However, these studies delve only into the types of outcomes pitchers might generate because of their repertoires, and not into the impacts of these pitch characteristics on either the actual offensive production that pitchers are tasked with suppressing or the monetary value of these skills within the baseball labor market. Ultimately, while these metrics have gained traction in baseball's public analytics community, their incorporation into meaningful formal analyses of pitchers' on-field skill and off-field compensation are sparse.

Because this grouping of metrics includes numbers that are immediately obvious and attractive to the naked eye of scouts and fans — especially when it comes to velocity and movement — it may be that certain characteristics directly related to pitchers' physical skills are being inappropriately valued by baseball front offices (as measured by player salaries) relative to the ability of those characteristics to contribute to offensive suppression and run prevention. By taking advantage of newly-available data provided by

baseball's tracking systems, a more thorough understanding of the player valuation process may be obtained, leading in turn to a more thorough understanding of the ability of baseball's front offices to properly value the big league pitchers on their rosters.

### **Empirical Framework**

In order to estimate the effect of pitch characteristics, including velocity, movement, and location, on a pitcher's ability to discourage offense at the MLB level, certain theoretical concepts and economic reasoning must be applied. First, we must establish an accurate measurement of offensive success against given pitch types to serve as a barometer for the performance pitchers are expected to prevent. This can be accomplished using a batter's weighted on-base average (wOBA). According to baseball website FanGraphs, wOBA "is a rate statistic which attempts to credit a hitter for the value of each outcome (single, double, etc.), rather than treating all hits or times on base equally." Developed by sabermetrician Tom Tango, wOBA essentially serves as a catch-all offensive statistic to appropriately weight the contributions of each at-bat outcome in terms of the value of that outcome towards ultimately scoring a run.

Each at-bat's wOBA, then, will serve as the dependent variable in the empirical analyses, which will consist of several ordered logit regressions. Within the scope of an individual at-bat, wOBA has six potential values, each of which is noted below with the on-field outcomes associated with them:

At-Bat Outcome	wOBA Value
Out	0.0
Walk	0.7
Single	0.9
Double	1.25
Triple	1.6
Home Run	2.0



The ordered logit model in its most general form allows us, then, to evaluate the independent impacts of various pitch characteristics on the probability of graduating to a higher wOBA level in an at-bat outcome. Essentially, it will assess the significance of pitch velocity, movement, and location on the likelihood of higher-quality at-bat outcomes. The research will focus mainly on these pitch characteristics, calculated by Major League Baseball's Statcast tracking systems and recorded on an at-bat-by-at-bat basis. Models will be estimated for each of the four primary pitch types in Major League Baseball — fastballs, sliders, curveballs, and change-ups.

Other factors will also be included among the estimated model's independent variables. These include release point on the horizontal and vertical axes, the relationship between batter and pitcher handedness, and the in-game circumstances under which an at-bat takes place (including balls, strikes, and outs).

Under this logic, the regression results are expected to provide three important insights into the ability of pitchers to prevent offensive production (i.e., to decrease the chances of higher-quality at-bat outcomes). First, velocity and vertical movement should be expected to have significant effects on pitching performance. Second, horizontal movement should not be expected to have a significant effect on pitching performance because, as much as a given pitch may break from east to west, horizontal movement suggests that the pitch still remains on the same plane as a hitter's bat, making it easier to hit. Third, the model can be expected to find that pitch location (as delineated by the organization of the strike zone in *Figure 1*) is a significant indicator of pitching performance at the Major League level.

Each of the dummy variables associated with a given pitch location is expected to have a negative sign attached to it once models are actually estimated. Because these variables will be in comparison to throwing a pitch over the middle of the plate, negative coefficients would suggest a decrease in offensive performance relative to throwing a pitch down the middle (where it is assumed it would be the easiest to hit). Similarly, because the wOBA dependent variable is a rate statistic where lower is better (from the perspective of a pitcher trying to prevent offense), negative coefficients on the other variables would suggest that as characteristics such as pitch velocity and movement improve (i.e., increase), a pitcher's performance would likewise improve, causing the

performance of the opposing offense to decrease. The outlier in terms of sign, *Balls*, is something one can expect to have a positive effect on the performance of opposing batters as they work their way into friendlier counts and are allowed more freedom to target specific pitches in specific zones.

Similarly, the second set of models within this analysis will use these same variables (averaged over full seasons) alongside player age as explanatory factors related to pitcher salary in the following season. In order for all of the salary data to operate on the same scale, each observation was adjusted into April 2017 dollars using Consumer Price Index information from the Bureau of Labor Statistics (Note: the month of April was elected instead of the more traditional January or December adjustments because of the typical calendar of a Major League Baseball season, which most often begins competitive play during the first week of April).

As such, the complete salary regression equation to be estimated, including the pitching statistic that will developed later in this paper, is described below:

$$\text{Adj\_Salary}_{it} = \alpha + \beta_1(\text{Age})_{it} + \beta_2(\text{Age\_Squared})_{it} + \beta_3(\text{ICD})_{it} + \beta_4(\text{vPitch})_{it} + \beta_5(\text{Pitch-X})_{it} + \beta_6(\text{Pitch-Z})_{it} + \beta_7Z_i + u_{it}$$

Based on the underlying theory of this analysis, it is expected that all of the coefficients in the model will be positive as a result of the underlying factors driving pitcher performance and therefore (hopefully) pitcher salary. In other words, it is unlikely that front offices are going so far as to decrease player salaries because of improvements in physical skill. However, it should also be expected that the *ICD* variable will be statistically insignificant, as command and to a lesser extent control are, as discussed earlier, less obvious skills to the naked eye and therefore may be undervalued or less important in determining player salaries. It should also be expected that the *Age Squared* variable is strongly significant, as players' wages can be expected to increase as they progress through their careers as a result of their eventual entrance into free agency.

## Data and Descriptive Statistics

In general, the data involved in this research project consists of at-bat information compiled through queries of the website Baseball Savant, which stores Statcast data recorded during Major League Baseball games. These statistics are presented largely as raw information (e.g., horizontal movement is broken down by inches, velocity is broken down by miles per hour, and so on). Each at-bat is treated as an independent and individual event, regardless of the season in which it took place or the participating players. The pitches included in the data set also each correspond to the end of an individual at-bat, meaning that pitches leading up to the outcome of a plate appearance (e.g., a fastball on an 0-0 count that a batter simply takes for a strike) are not included in the data.

Thorough descriptions of the variables listed above and their sources can be found in *Tables 1* and *3* below, while descriptive statistics of all of these variables can be found in *Tables 2* and *4*. Attention should be given to the range of the *Adjusted Salary* variable, which stretches from the adjusted major league minimum in 2007 (\$341,068) to the maximum of \$35,571,429 belonging to Clayton Kershaw of the Los Angeles Dodgers in 2017. The wide range of potential payouts to pitchers specifically speaks to the significant possibility of over- or under-spending by major league teams on their pitching staffs, which in turn opens up the potential for inefficiencies in the use of team payrolls that form the basis for this paper.

An analysis of the correlation between variables presents interesting insights as well. The velocity of a pitcher's fastball shows a strong correlation with the velocity of his other pitches, something that makes sense in theory and may require the inclusion of just one or two velocity variables. Similarly, the movement of certain types of similar pitches (i.e., fastballs and change-ups, curveballs and sliders) show strong correlations and may require a similar adjustment. The most obvious of these (and therefore the most threatening to the following analyses) is the correlation in horizontal movement between fastballs and changeups, which can be explained by their similar release and axis tilt

coming out of the pitcher's hand. This problem is counteracted by excluding the *Changeup Horizontal* variable from the regression models.

Ultimately though, based on the breadth of available information in the data and lack of correlation between the vast majority of the variables, the data appears to be more than capable of providing some answers when it comes to the effect of location, velocity, and movement on pitching performance in baseball.

## **Empirical Results**

In order to answer the research question at hand, we must analyze the impact of various components of pitch physics on offensive outcomes in Major League Baseball. This will be done through the aforementioned ordered logit models. Robust standard errors were used to correct for heteroskedasticity identified through statistical testing across the relevant pitch samples. Special attention should be given to the impact of the *Pitch Location* variables as an indicator of the impact of pitch command on offensive outcomes.

The complete results for all four pitch types can be found in *Tables 5-8* below.

Overall, location proves to be statistically significant in the vast majority of Statcast zones for all four pitches. This means that, relative to locating a pitch over the middle of the plate, locating a pitch in another part of the strike zone has a statistically significant impact on offensive performance. As expected, the vast majority of these impacts are also negative, and most are statistically significant at the 1 percent level. From the perspective of the pitcher, this is a good thing, as lower offensive performance indicates improved ability in suppressing offense by the pitcher.

Because of the comparison of each pitch location zone to throwing a pitch down the middle (where it should be assumed to be the easiest to hit), results that produce positive coefficients or are not statistically significant also are able to offer interesting insights. The most obvious of these are in relation to fastballs, where pitches thrown in Zones 4, 8, and 11 are not statistically significant in suppressing offense (relative to pitches thrown in zone 5) and pitches thrown in zones 13 and 14 actually improve the likelihood of higher quality hits. While the interpretation of the non-significant zones

simply means that there is no discernable difference in offensive performance against these pitches relative to a pitch thrown down the middle (statistically, they are just as easy to produce quality contact against), the interpretation of the positive coefficients requires some basic baseball intuition. It is likely the case that fastballs starting below the strike zone are easier for hitters to take, meaning it is likelier for pitches in these zones to more consistently result in walks (which are included in the range of wOBA outcomes). It is assumed, then, that the higher likelihood of quality offensive performance results from this combination of free-base tendencies in addition to the outcomes of pitches that are actually put into play.

The incorporation of other pitch and game-context characteristics within the models also yield several interesting takeaways. Velocity proves to be statistically significant for fastballs, curveballs, and sliders, meaning that the importance of pitch location in determining the likelihood of improved offensive performance should not overshadow the equal importance of velocity in these same determinations. Vertical release point also proves significant for all four pitch types, while vertical movement is significant for fastballs, sliders, and change-ups. As expected, horizontal movement is not significant for any pitch type, and it is assumed that, as was mentioned earlier, this is indeed caused by the planar relationship between ball and bat. Finally, ball and strike counts are statistically significant for each pitch type, as they factor heavily into the approach of the opposing hitter on a pitch-by-pitch basis. As an example, a curveball on a 3-0 count is unlikely to even generate a swing (with the batter factoring in the probability of a walk), while a hitter facing a curveball with two strikes is far more likely to be encouraged to be defensive and simply do what he can with the pitch to avoid a strikeout.

As a result, it can be demonstrated that pitch physics, including velocity and movement, have a statistically significant impact on a given pitcher's ability to prevent better offensive performances at the Major League level. Specific to the research question at hand, a pitcher's ability to locate the ball in certain parts of the strike zone — and at times intentionally outside of the strike zone — has a corresponding impact on the types of hit quality that pitcher will allow. This suggests that the development of a statistic to evaluate the pitchers who throw their pitches in these zones most consistently would improve the ability of MLB front offices to objectively evaluate pitchers alongside

velocity and stuff (i.e., movement). This could in turn improve on the efficiency of MLB's collective spending on pitchers.

### **Development of the Indexed Command Diagnostic**

The public baseball analytics community has produced several attempts to measure pitch command as a statistic. In 2010, Sportsvision and MLB Advanced Media released COMMANDf/x, which uses radar tracking systems to determine the distance between the position of the catcher's glove when the pitch is released (i.e., where the pitcher is likely to be aiming) and the pitch's actual location. However, it was quickly found that differences in how catchers set up prior to receiving pitches had far too heavy of an impact on the COMMANDf/x outputs of the pitchers they were catching for. This was especially true for breaking balls, meaning that COMMANDf/x was only truly useful for evaluating fastball command if it were useful at all.

Others used the final pitch locations themselves to evaluate pitch command. Petti (2013) developed Edge Percentage (Edge%) and its corresponding brother-stat, Heart Percentage (Heart%), to evaluate the pitchers who best locate their pitches around the edges of the strike zone (presumably in the locations that are hardest for opposing batters to hit) as well as best avoid throwing pitches over the middle (i.e., "heart") of the plate. This encouraged Brooks, Pavlidis, and Judge (2015) to unveil a statistic termed Called Strikes Above Average (CSAA), which combined the contributions of the catcher, pitcher, and umpire into a mixed-model approach. Originally used as a way of measuring pitch framing among catchers (another skill that the industry has begun to emphasize in the past few seasons), the model can also be reworked to evaluate pitchers' ability to locate pitches in zones where they are more likely to generate called strikes from the umpire. Both Edge% and CSAA, however, don't factor in the benefit of different pitch locations to the performance of the pitcher relative to other locations.

The best advancement in measuring pitch command, as it turns out, is also the most recent. The statistic, entitled Command+ and developed by Sarris (2018) in conjunction with analysts at STATS LLC, incorporates pitch intent in the evaluation of command. This means that a back foot curveball that dives out of the bottom of the strike

zone and towards a hitter's feet can be considered as positively as a fastball that is spotted on the outside corner. While an improvement on other prior pitch command stats, where Command+ falls short is in the subjective determinations of what a pitch should do and the corresponding impact of quality movement on the evaluation of pitch command. As a result, Command+ trends closer to evaluating the collective quality of a pitch instead of isolating pitch command as an individual and independent pitching skill.

As previously mentioned, the results of the ordered logit regressions analyzing the impact of pitch command on MLB at-bat outcomes can in turn be used to develop a statistic that improves on these prior attempts to quantify pitch command as a skill. This can be accomplished using the following (and confidently safe) assumption: armed with the knowledge that pitch location does play a role in the performance of opposing batters, pitchers can be expected to want to throw pitches as often as possible in locations where they are least likely to get hit hard. With this intent in mind, we can then measure the frequency with which pitchers throw their pitches into each zone to evaluate the success in accomplishing this intention.

The calculation of the statistic that results, labeled as a pitcher's *Indexed Command Diagnostic* by the author (and henceforth referred to as ICD), incorporates weights based on the average offensive performance against pitches in each zone, filtered by pitch type. These weights accomplish the goal of giving pitchers credit for throwing pitches in locations where those pitches will be harder to hit, and penalizing pitchers for throwing pitches in locations where those pitches are more likely to be hit hard. By combining these weights (outlined more thoroughly for the 2017 season in *Figures 2-5* below) with the frequency with which a given pitcher throws the ball in each spot, ICD measures the percentage of offense a pitcher can be expected to prevent or allow via command of a given pitch type. ICD can be divided by pitch type or combined into an overall pitch command statistic, where the summation of ICD measurements for each pitch type into one collective statistic would reward pitchers who are able to command multiple types of pitches well.

The resulting rankings of pitchers with the best and worst command of their pitches for 2017 are outlined in *Tables 9 and 10* below.

## Validity of the Indexed Command Diagnostic

While it is fortunate that ICD seems to cover many of the theoretical weaknesses of other attempts to quantify pitch command, it would be irresponsible to claim that a mathematical combination of a handful of data points could immediately be considered a useful evaluative tool. In order to truly determine the validity of ICD as a pitching statistic, a couple of statistical evaluations of the metric must be performed.

First, the year-to-year correlations of ICD, both by pitch type and in total, should indicate whether or not ICD accurately and appropriately measures a true skill. If it is to be assumed that pitch command is a skill that certain pitchers inherently possess (and others inherently lack), the correlation between ICD production in one season and ICD production in the next should be relatively strong.

The year-to-year correlations between ICD production in 2016 and 2017 (the two seasons for which fully available Statcast data on pitch location exists) are outlined below:

Fastball ICD: 0.4809

Curveball ICD: 0.6515

Slider ICD: 0.6675

Change-up ICD: 0.3261

Total ICD: 0.5379

Outside of change-ups, the ICD correlations between seasons are encouraging, and serve to validate the fact that ICD as a statistic is measuring a real skill that carries over from one season to the next.

Second, the statistical relationships between ICD and other pitching stats can be used to evaluate the benefits of ICD in measuring a skill that actually contributes to on-field success. If we are to assume that ICD measures pitch command and that better pitch command is associated with improved offensive suppression (and therefore improved run prevention) by a given pitcher, then ICD should have a meaningful relationship with



widely accepted run prevention statistics such as Earned Run Average (ERA) and Fielding Independent Pitching (FIP).

These results are outlined in full in *Tables 11 and 12* below. In summation, however, the results are again promising in indicating the benefits of ICD as a legitimate statistic. Both Fastball ICD and Slider ICD are statistically significant in explaining variation in ERA between pitchers. Meanwhile, FIP — a run-prevention statistic that attempts to isolate contributions made by the pitcher from those made by the fielders behind him in keeping runs off the scoreboard — demonstrates even more encouraging results, with Fastball ICD, Slider ICD, and Change-up ICD all being statistically significant in explaining variation in pitcher FIP. The incorporation of these ICD statistics with other pitch characteristics produces a FIP model with an adjusted R-squared value of  $R^2 = 0.4791$ .

Ultimately, both analyses demonstrate that not only is ICD theoretically sound in incorporating appropriate factors into its calculation, but it also carries an encouraging amount of statistical weight in terms of evaluating on-field skills inherent to individual pitchers.

### **Financial Value of Pitch Command**

With the confirmation that ICD is in fact a meaningful pitching statistic, the final step in determining how much its incorporation into the industry would impact the efficiency of MLB pitcher contracts involves an analysis of ICD's relationship with player salaries. While of course ICD in and of itself has not been factored into the recent contractual decisions of MLB front offices (the statistic did not even exist more than a handful of pages ago), if pitch command is a factor in determining player salaries, it is more likely that ICD would be able to have a meaningful impact on the financial incentives surrounding MLB pitchers.

This relationship between ICD in one season and player salaries in the following season can be assessed using ordinary least squares (OLS) regressions. The results of these stepwise regressions can be found in *Table 13* below. In short, though, the data demonstrates no statistically significant relationship between ICD and player salary,

indicating that it is an undervalued skill relative to its on-field impact in preventing offense.

However, this also holds true for other pitch characteristics, including pitch velocity, horizontal movement, and vertical movement. This begs the question: if none of the skills inherent to pitchers and their repertoires carry monetary value under baseball's current financial systems, what are front offices paying for when they evaluate pitchers? The answer, as it turns out, is statistics rather than skills. When examining similar salary models with pitching statistics as explanatory variables, several carry significance in determining variation in player salaries. These results, which can be found in *Table 14*, indicate the statistical significance of age, strikeout rate, walk rate, and home run rate within the model.

Age — both *Age* and *Age Squared*, in the model — is an obvious inclusion: as players get older and their careers progress, they become eligible for league-mandated raises on their rookie paychecks (which are often simply the league minimum) in the form of arbitration eligibility and, eventually, free agency.

The other variables, however, suggest that front offices are far more willing to pay for actual on-field production instead of the underlying skills that actually drive this production. This makes sense in some aspects: games are won and lost on the same fields where these results ultimately occur, and the infancy of the availability of data on pitch characteristics may preclude the inclusion of this data in players' salaries. However, financial ignorance towards these underlying skills carries the potential for dangerous regressions in performance and allows for the greater possibility of luck playing a role in deciding these outcomes and, ultimately, the monetary rewards they fuel. By focusing instead on inherent pitching skills such as pitch velocity, movement, and command, it is likelier in theory that some of these drop-offs in performance could be identified ahead of financial commitments to players, thereby offsetting much of the risk associated with the amount of money handed out to MLB arms on an annual basis.

## Conclusion

Ultimately, the research suggests takeaways similar to those hypothesized prior to the aforementioned empirical analyses. While velocity and movement do have significant impacts at times, that effect is limited to certain pitch types and is more beneficial when pitches move north to south rather than east to west. Furthermore, the independent effects of pitch physics cannot be fully understood without the corresponding incorporation of pitch command, a skill that is undervalued by front office executives at the major league level and measurable via the development of the Indexed Command Diagnostic. These findings not only agree with the original intentions of this research, but also serve to reemphasize and reward the ability to throw the ball where one intends in an era of higher velocities and exaggerated arm injury rates.

This is not to suggest that ICD as it currently exists is a perfect statistic. There are meaningful drawbacks that will need to be ironed out before the statistic may be implemented to its fullest potential. The most obvious of these is the selection bias of the outcome-oriented data that makes up the analyses included in this paper. True pitch-by-pitch data is not available publicly, and as such many thousands of well-placed pitches were not included in developing the statistic and using it to evaluate MLB pitchers. By including all pitches thrown by MLB pitchers, a more complete picture of each pitcher's skill in locating his pitches could be developed while eliminating some of the small samples of pitch types that may have impacted individual ICD scores for some players.

One must also consider the potential mischaracterization of pitch types by the Statcast algorithms that actually label each pitch. At times, Statcast did not seem to recognize pitchers as throwing certain pitch types despite these pitchers having well-known offerings of those excluded pitch types. For example, Statcast does not acknowledge Boston's Drew Pomeranz as throwing a curveball despite his curveball being one of the most well-known breaking balls in the league. While this could obviously be fixed via refinement of the Statcast algorithms, this issue could also be overcome through a wider segmentation of pitch types in various classes (e.g., include all breaking balls rather than only curveballs and sliders, and therefore evaluate a pitcher's

total breaking ball command), and would again help in raising the thresholds for pitchers to have enough observations of a given pitch type to make their ICD outputs meaningful.

Finally, it should be mentioned that ICD ignores the impacts of individual differences between hitters, as well as the contributions sequencing (i.e., the order in which pitches are thrown during an at-bat) may have on pitch effectiveness. However, the incorporation of these would not only require their calculation (a task that has proven difficult for baseball analysts in the past), but would also significantly raise the computational costs of determining ICD. Further honing of the statistic through the processes recommended above may also reconcile some of these analytical challenges.

In the end, though, these improvements would ultimately serve as mere refinements to what is already a useful addition to the pitching statistics library. Not only is it clear that pitch location has an impact on offensive outcomes (and therefore pitcher performance) and lacks the financial clout it deserves within the baseball labor market, but the implementation of ICD as a tool to evaluate pitchers is a statistically meaningful method of furthering league emphasis on pitch command and its associated impact on pitcher longevity over the course of big league careers. The development of ICD, then, can be seen as a significant step closer to a championship-winning competitive advantage for front offices — allowing MLB executives to more efficiently gauge appropriate contractual rewards — while also incentivizing more productive pitcher behavior on the field, prolonging careers while avoiding sacrifices to the run prevention objective that forms the foundation of a pitcher's livelihood.

## References

- Bradbury, J.C. (2007). Does the Baseball Labor Market Properly Value Pitchers? *Journal of Sports Economics*, 8(6), 616-632.
- Brooks, D., Pavlidis, H., & Judge, J. (2015, February 5). Moving beyond WOWY: A Mixed Approach to Measuring Catcher Framing. Baseball Prospectus. Retrieved May 6, 2018, from <https://www.baseballprospectus.com/news/article/25514/moving-beyond-wowy-a-mixed-approach-to-measuring-catcher-framing/>
- FOX Sports (2015, February 2). What makes a great curveball? FOX Sports. Retrieved March 24, 2018, from <https://www.foxsports.com/mlb/just-a-bit-outside/story/best-curveball-corey-kluber-felix-hernandez-craig-kimbrel-jose-fernandez-020215>
- Gonzalez, R. (2018, April 12). Why it's almost impossible for fastballs to get any faster. Wired. Retrieved May 2, 2018, from <https://www.wired.com/story/why-its-almost-impossible-for-fastballs-to-get-any-faster/>
- Hakes, J. K., & Sauer, R. D. (2006). An economic evaluation of the Moneyball hypothesis. *Journal of Economic Perspectives*, 20, 173-185.
- Kahn, L. M. (1993). Free agency, long-term contracts and compensation in Major League Baseball: Estimates from panel data. *The Review of Economics and Statistics*, 75, 157-164.
- Krautmann, A. C. (1999). What's wrong with Scully-estimates of a player's marginal revenue product. *Economic Inquiry*, 37, 369-381.
- Lewis, M. L. (2003). *Moneyball: The art of winning an unfair game*. New York: Norton.
- Lindbergh, B. (2014, July 24). Don't call it a Lincecum comeback. Grantland. Retrieved May 6, 2018 from <http://grantland.com/features/tim-lincecum-rebound-san-francisco-giants/>
- McCracken, V. (2001, January 23). Pitching and defense: How much control do hurlers have? Baseball Prospectus. Retrieved October 21, 2017, from <http://www.baseballprospectus.com/article.php?articleid=878>
- Miller, S. (2015, May 19). Pitch 22. ESPN. Retrieved October 22, 2017 from [http://www.espn.com/mlb/story/\\_/id/12842678/why-tommy-john-surgeries-cease-soon](http://www.espn.com/mlb/story/_/id/12842678/why-tommy-john-surgeries-cease-soon)
- Pavlidis, H. (2013, August 30). What makes a good change-up: An investigation, part 3. Baseball Prospectus. Retrieved March 24, 2018, from <https://www.baseballprospectus.com/news/article/21675/what-makes-a-good-changeup-an-investigation-part-three/>

Pemstein, J. (2015, April 29). Batted balls: It's all about location, location, location. FanGraphs. Retrieved October 22, 2017 from <https://www.fangraphs.com/blogs/batted-balls-its-all-about-location-location-location/>

Pemstein, J. (2015, June 3). On rotation, part 2: The effects of spin on pitch outcomes. FanGraphs. Retrieved March 24, 2018 from <https://www.fangraphs.com/blogs/on-rotation-part-2-the-effects-of-spin-on-pitch-outcomes/>

Petti, B. (2013, January 15). The difference pitching on the edge makes. FanGraphs. Retrieved May 6, 2018, from <https://www.fangraphs.com/blogs/the-difference-pitching-on-the-edge-makes/>

Sarris, E. (2015, December 16). A different way to look at sliders. FanGraphs. Retrieved March 24, 2018, from <https://www.fangraphs.com/blogs/a-different-way-to-look-at-sliders/>

Sarris, E. (2017, April 25). Why we still don't have a great command metric. FanGraphs. Retrieved May 6, 2018, from <https://www.fangraphs.com/blogs/why-we-still-dont-have-a-great-command-metric/>

Sarris, E. (2018, May 10). Exclusive: A big step forward in measuring command. The Athletic. Retrieved May 11, 2018, from <https://theathletic.com/346863/2018/05/10/exclusive-a-big-step-forward-in-measuring-command/>

Scully, G. W. (1974). Pay and performance in Major League Baseball. *American Economic Review*, 64, 915-930.

Scully, G. W. (1989). *The business of Major League Baseball*. Chicago: University of Chicago Press.

Zimbalist, A. S. (1992). *Baseball and billions*. New York: Basic Books.

## Appendix

Figure 1: Statcast Delineation of the Strike Zone

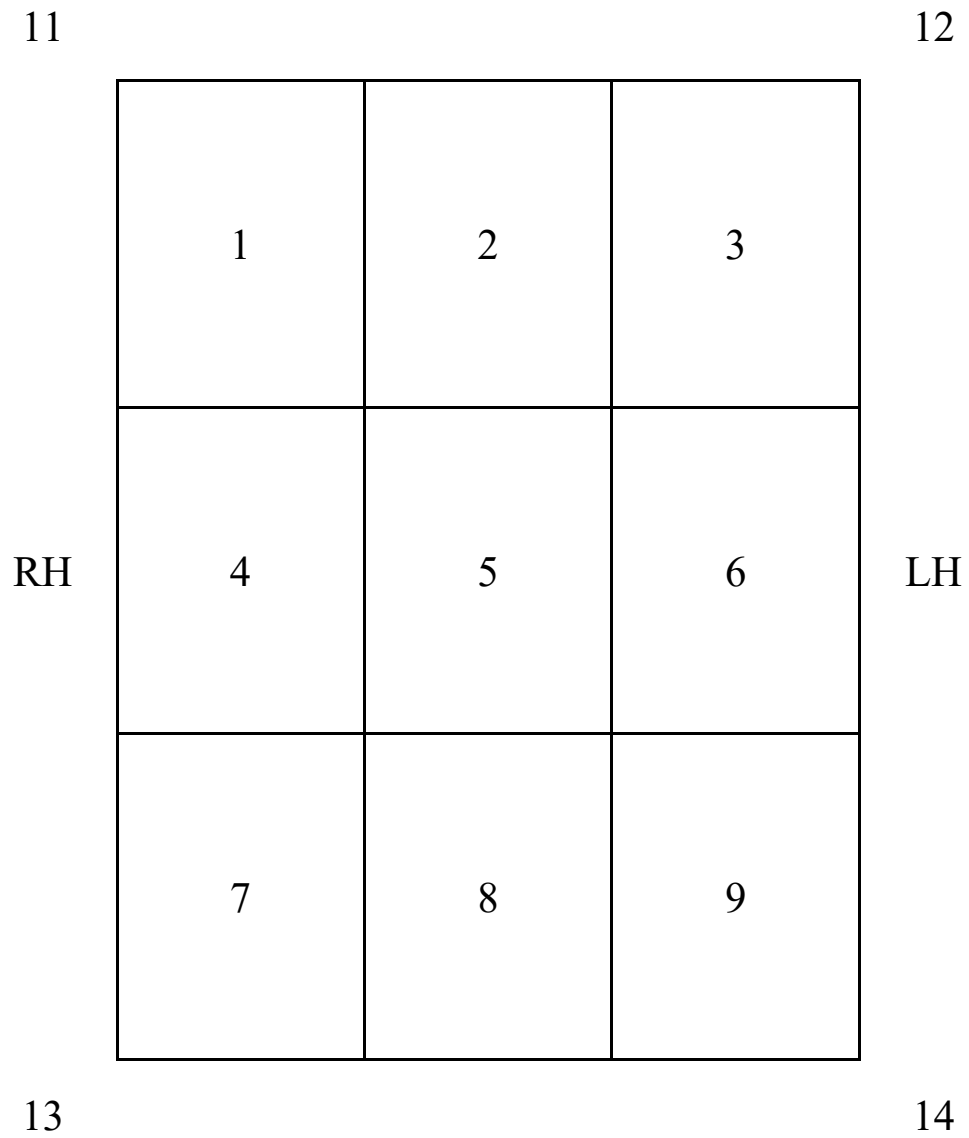




Figure 2: Fastball Weights for ICD Formula (2017 Season)

	1.87			14.09
		14.56	2.41	16.65
RH		-1.62	-12.75	2.15
		10.32	-4.91	11.65
	-26.03			-28.43

Figure 3: Curveball Weights for ICD Formula (2017 Season)

-51.73

-34.80

	-4.51	-2.69	13.11	
RH	-28.40	-50.58	-8.98	LH
	-2.19	-7.39	10.60	

26.04

40.33

Figure 4: Slider Weights for ICD Formula (2017 Season)

-29.86

3.30

	11.37	-19.94	6.37	
RH	-27.12	-47.39	-14.59	LH
	-9.15	-21.02	11.06	

27.71

36.61

Figure 5: Change-up Weights for ICD Formula (2017 Season)

-22.90

-4.26

	-23.25	-47.97	-12.52	
RH	-15.51	-54.29	-14.32	LH
	18.10	-1.63	9.96	

28.50

25.56

Table 1: Data Definitions and Sources (Pitch-by-Pitch Data)

Variable	Definition
wOBA*	Offensive statistic that measures the contributions of various types of at-bat outcomes towards scoring a run
Zone (#)*	Location of the recorded pitch as measured by MLB's Statcast systems and delineated in <i>Figure 1</i>
Velocity*	Velocity of the recorded pitch as measured in miles per hour
Horizontal Release Point*	Horizontal release point of the recorded pitch, as measured by feet from the center of the pitching rubber
Vertical Release Point*	Vertical release point of the recorded pitch, as measured by feet from the ground
Same Hand*	Dummy variable that evaluates whether the pitcher and batter are performing their tasks with the same hand (Same=1)
Horizontal Pitch Movement*	Horizontal movement of the given pitch as measured by inches moved from the catcher's left to the catcher's right, relative to that pitch remaining straight upon release
Vertical Pitch Movement*	Vertical movement of the pitch as measured by inches moved from the bottom of the strike zone (from the catcher's perspective) to the top, relative to that pitch remaining straight upon release
Outs*	Number of outs when recorded pitch occurred
Balls*	Number of balls when recorded pitch occurred
Strikes*	Number of strikes when recorded pitch occurred

\*Data courtesy of Baseball Savant

Table 2: Descriptive Statistics (Pitch-by-Pitch Data)

Variable	Mean	St. Dev.	Minimum	Maximum
Fastball				
Velocity	93.307	2.794	61.200	105.400
Horizontal Movement	-0.340	0.728	-2.403	2.364
Vertical Movement	1.443	0.316	-0.340	3.169
Horizontal Release	-0.836	1.844	-4.992	4.477
Vertical Release	5.974	0.429	3.262	7.466
Curveball				
Velocity	78.506	4.083	45.100	97.400
Horizontal Movement	0.357	0.770	-2.347	2.374
Vertical Movement	-0.566	0.514	-2.254	3.818
Horizontal Release	-0.876	1.838	-5.019	5.142
Vertical Release	5.955	0.462	2.929	7.361
Slider				
Velocity	84.942	3.344	59.400	97.900
Horizontal Movement	0.158	0.537	-2.908	2.522
Vertical Movement	0.331	0.452	-1.563	2.787
Horizontal Release	-1.067	1.892	-5.341	5.728
Vertical Release	5.894	0.508	0.007	8.608
Change-up				
Velocity	84.141	3.580	54.900	97.000
Horizontal Movement	-0.389	1.089	-3.577	2.391
Vertical Movement	0.786	0.465	-1.480	2.633
Horizontal Release	-0.620	2.033	-4.628	4.962
Vertical Release	5.880	0.479	0.083	7.567

Table 3: Data Definitions and Sources (Seasonal Data)

Variable	Definition
FIP*	Run-prevention metric that utilizes only defense-independent statistics in its calculation
Adjusted Salary**	Measure of a player's value from the perspective of MLB front offices (adjusted into April '17 dollars)
Age*	Nominal value that serves as a proxy for a player's experience in the league
(Pitch) ICD	New statistic that describes the contributions of pitch command toward preventing offensive performance
(Pitch) Velocity*	Velocity of the given pitch as measured in miles per hour
(Pitch) Horizontal*	Horizontal movement of the given pitch as measured by inches moved from the catcher's left to the catcher's right, relative to that pitch remaining straight upon release
(Pitch) Vertical*	Vertical movement of the pitch as measured by inches moved from the bottom of the strike zone (from the catcher's perspective) to the top, relative to that pitch remaining straight upon release

---

\*Data courtesy of FanGraphs

\*\*Data courtesy of Baseball Prospectus and FOX Sports

Table 4: Descriptive Statistics (Seasonal Data)

Variable	Mean	St. Dev.	Minimum	Maximum
FIP	4.184	0.730	1.800	6.450
ERA	4.124	0.891	1.690	6.810
Age	28.608	3.843	22.000	44.000
ICD	-0.774	9.281	-26.498	29.474
Fastball Velocity	92.115	2.395	82.300	97.900
Slider Velocity	84.398	2.906	75.700	90.800
Curveball Velocity	77.296	3.368	67.100	84.800
Change-up Velocity	84.443	2.769	71.300	90.800
Fastball Horizontal	-1.900	4.996	-9.700	11.300
Slider Horizontal	1.159	2.725	-7.500	8.300
Curveball Horizontal	3.084	4.187	-8.500	10.300
Change-up Horizontal	-2.540	7.450	-11.300	12.200
Fastball Vertical	9.315	1.521	3.400	13.500
Slider Vertical	2.307	2.271	-4.200	9.200
Curveball Vertical	-4.676	2.579	-11.300	2.300
Change-up Vertical	5.070	2.180	-0.100	12.700
Adj. Salary	7,195,741.000	8,153,450.000	374,144.000	35,600,000.000



Table 5: Ordered Logit Results (Fastballs)

VARIABLES	(1) wOBA	(2) wOBA	(3) wOBA	(4) wOBA	(5) wOBA
Zone 1	-0.356*** [0.035]	-0.342*** [0.033]	-0.342*** [0.033]	-0.330*** [0.030]	-0.322*** [0.030]
Zone 2	-0.210*** [0.033]	-0.196*** [0.030]	-0.196*** [0.030]	-0.184*** [0.028]	-0.176*** [0.028]
Zone 3	-0.288*** [0.035]	-0.274*** [0.033]	-0.274*** [0.033]	-0.263*** [0.031]	-0.254*** [0.030]
Zone 4	-0.060** [0.030]	-0.047* [0.028]	-0.047* [0.028]	-0.036 [0.025]	
Zone 6	-0.131*** [0.031]	-0.117*** [0.028]	-0.117*** [0.028]	-0.106*** [0.026]	-0.097*** [0.025]
Zone 7	-0.113*** [0.036]	-0.100*** [0.033]	-0.100*** [0.033]	-0.088*** [0.031]	-0.080*** [0.031]
Zone 8	-0.038 [0.033]				
Zone 9	-0.187*** [0.035]	-0.173*** [0.033]	-0.173*** [0.033]	-0.161*** [0.030]	-0.153*** [0.030]
Zone 11	-0.045* [0.024]	-0.029 [0.021]	-0.029 [0.021]		
Zone 12	-0.086*** [0.026]	-0.076*** [0.023]	-0.076*** [0.023]	-0.063*** [0.019]	-0.055*** [0.019]
Zone 13	0.101*** [0.026]	0.113*** [0.023]	0.113*** [0.023]	0.126*** [0.019]	0.134*** [0.018]
Zone 14	0.186*** [0.025]	0.201*** [0.021]	0.201*** [0.021]	0.214*** [0.017]	0.221*** [0.017]
Velocity	-0.014*** [0.002]	-0.013*** [0.002]	-0.013*** [0.002]	-0.013*** [0.002]	-0.013*** [0.002]
Vertical Release	0.099*** [0.014]	0.096*** [0.014]	0.096*** [0.014]	0.096*** [0.014]	0.096*** [0.014]
Vertical Movement	-0.204*** [0.019]	-0.203*** [0.019]	-0.203*** [0.019]	-0.204*** [0.019]	-0.204*** [0.019]
Outs	-0.035*** [0.007]	-0.035*** [0.007]	-0.035*** [0.007]	-0.035*** [0.007]	-0.035*** [0.007]
Balls	0.273*** [0.006]	0.272*** [0.005]	0.272*** [0.005]	0.271*** [0.005]	0.272*** [0.005]
Strikes	-0.456*** [0.007]	-0.457*** [0.007]	-0.457*** [0.007]	-0.457*** [0.007]	-0.457*** [0.007]
<b>Observations</b>	<b>124,991</b>	<b>124,991</b>	<b>124,991</b>	<b>124,991</b>	<b>124,991</b>

Robust standard errors in brackets

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

Table 6: Ordered Logit Results (Curveballs)

VARIABLES	(1) wOBA	(2) wOBA	(3) wOBA
Zone 1	-0.432*** [0.097]	-0.447*** [0.091]	-0.448*** [0.091]
Zone 2	-0.323*** [0.090]	-0.338*** [0.085]	-0.338*** [0.085]
Zone 3	-0.496*** [0.117]	-0.512*** [0.112]	-0.513*** [0.112]
Zone 4	-0.166** [0.068]	-0.179*** [0.061]	-0.179*** [0.061]
Zone 6	-0.298*** [0.071]	-0.315*** [0.064]	-0.317*** [0.064]
Zone 7	-0.255*** [0.069]	-0.267*** [0.061]	-0.267*** [0.061]
Zone 8	-0.218*** [0.060]	-0.232*** [0.052]	-0.232*** [0.052]
Zone 9	-0.396*** [0.066]	-0.413*** [0.058]	-0.414*** [0.058]
Zone 11	0.054 [0.062]		
Zone 12	0.009 [0.078]		
Zone 13	-0.538*** [0.058]	-0.551*** [0.047]	-0.549*** [0.047]
Zone 14	-0.712*** [0.054]	-0.729*** [0.043]	-0.730*** [0.043]
Velocity	-0.010*** [0.004]	-0.010*** [0.004]	-0.011*** [0.003]
Horizontal Release	-0.021* [0.012]	-0.021* [0.012]	-0.015* [0.008]
Vertical Release	0.056* [0.031]	0.057* [0.031]	0.061** [0.030]
Vertical Movement	-0.013 [0.029]	-0.013 [0.029]	
Balls	0.313*** [0.015]	0.315*** [0.015]	0.316*** [0.015]
Strikes	-0.649*** [0.022]	-0.649*** [0.022]	-0.649*** [0.022]
Outs	0.016 [0.017]	0.016 [0.017]	
<b>Observations</b>	<b>28,992</b>	<b>28,992</b>	<b>28,992</b>

Robust standard errors in brackets

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

Table 7: Ordered Logit Results (Sliders)

VARIABLES	(1) wOBA	(2) wOBA
Zone 1	-0.567*** [0.075]	-0.566*** [0.075]
Zone 2	-0.332*** [0.065]	-0.332*** [0.065]
Zone 3	-0.545*** [0.071]	-0.545*** [0.071]
Zone 4	-0.184*** [0.052]	-0.186*** [0.052]
Zone 6	-0.223*** [0.045]	-0.222*** [0.045]
Zone 7	-0.186*** [0.052]	-0.191*** [0.052]
Zone 8	-0.097** [0.044]	-0.097** [0.044]
Zone 9	-0.458*** [0.045]	-0.455*** [0.045]
Zone 11	-0.099** [0.045]	-0.095** [0.045]
Zone 12	-0.273*** [0.052]	-0.276*** [0.052]
Zone 13	-0.540*** [0.043]	-0.551*** [0.042]
Zone 14	-0.615*** [0.037]	-0.612*** [0.037]
Velocity	-0.018*** [0.003]	-0.017*** [0.003]
Vertical Release	0.083*** [0.019]	0.079*** [0.019]
Same Hand	-0.069*** [0.020]	-0.065*** [0.020]
Balls	0.329*** [0.010]	0.330*** [0.010]
Strikes	-0.577*** [0.015]	-0.579*** [0.015]
Vertical Movement	0.050** [0.023]	0.048** [0.023]
Outs	-0.032*** [0.012]	-0.032*** [0.012]
<b>Observations</b>	<b>59,238</b>	<b>59,238</b>

Robust standard errors in brackets

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

Table 8: Ordered Logit Results (Change-ups)

VARIABLES	(1) wOBA	(2) wOBA	(3) wOBA
Zone 1	-0.346*** [0.086]	-0.353*** [0.086]	-0.354*** [0.086]
Zone 2	-0.168* [0.096]	-0.166* [0.096]	-0.163* [0.096]
Zone 3	-0.478*** [0.107]	-0.463*** [0.106]	-0.466*** [0.106]
Zone 4	-0.300*** [0.060]	-0.311*** [0.059]	-0.314*** [0.059]
Zone 6	-0.424*** [0.066]	-0.410*** [0.065]	-0.413*** [0.065]
Zone 7	-0.520*** [0.059]	-0.530*** [0.058]	-0.536*** [0.058]
Zone 8	-0.391*** [0.056]	-0.391*** [0.056]	-0.393*** [0.056]
Zone 9	-0.591*** [0.061]	-0.580*** [0.061]	-0.582*** [0.061]
Zone 11	-0.164*** [0.055]	-0.175*** [0.054]	-0.178*** [0.054]
Zone 12	-0.354*** [0.065]	-0.333*** [0.064]	-0.340*** [0.063]
Zone 13	-0.636*** [0.050]	-0.645*** [0.050]	-0.651*** [0.049]
Zone 14	-0.587*** [0.050]	-0.580*** [0.050]	-0.583*** [0.050]
Velocity	-0.001 [0.003]		
Vertical Release	-0.096*** [0.026]	-0.096*** [0.026]	-0.095*** [0.026]
Same Hand	0.047* [0.027]	0.037 [0.026]	
Balls	0.303*** [0.011]	0.301*** [0.011]	0.301*** [0.011]
Strikes	-0.495*** [0.016]	-0.494*** [0.016]	-0.491*** [0.016]
Vertical Movement	0.090*** [0.027]	0.093*** [0.027]	0.092*** [0.027]
Outs	-0.015 [0.014]		
<b>Observations</b>	<b>39,344</b>	<b>39,344</b>	<b>39,344</b>

Robust standard errors in brackets

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

Table 9: Top of ICD Rankings, 2017 Season

Player	Team	ICD
Mike Clevinger	Indians	28.99
Zack Greinke	Diamondbacks	26.80
Jon Lester	Cubs	23.38
Kyle Gibson	Twins	20.58
Carlos Carrasco	Indians	18.15
Tyler Chatwood	Rockies	14.17
Bartolo Colon	Braves	12.87
Dallas Keuchel	Astros	12.37
Zack Godley	Diamondbacks	12.36
J.A. Happ	Blue Jays	11.56
Alex Wood	Dodgers	11.56
Tanner Roark	Nationals	11.45
Kyle Freeland	Rockies	10.83
Aaron Nola	Phillies	10.79
Johnny Cueto	Giants	10.74
Sonny Gray	Athletics	10.50
Hyun-Jin Ryu	Dodgers	9.78
Marcus Stroman	Blue Jays	9.39
Eduardo Rodriguez	Red Sox	9.03
Wade Miley	Orioles	8.39

\*Minimum of 120 IP and 10 pitches per type to qualify

Table 10: Bottom of ICD Rankings, 2017 Season

Player	Team	ICD
German Marquez	Rockies	-26.50
James Paxton	Mariners	-22.73
Ubaldo Jimenez	Orioles	-22.34
Andrew Cashner	Rangers	-16.76
Jhoulys Chacin	Padres	-15.48
Matt Cain	Giants	-15.43
Parker Bridwell	Angels	-14.71
Michael Wacha	Cardinals	-14.60
Antonio Senzatela	Rockies	-14.50
Chad Kuhl	Pirates	-12.34
Rick Porcello	Red Sox	-12.06
Jason Hammel	Royals	-11.96
Jose Urena	Marlins	-10.95
Jharel Cotton	Athletics	-10.62
Brad Peacock	Astros	-10.01
Jordan Zimmermann	Tigers	-9.66
Miguel Gonzalez	White Sox	-9.23
JC Ramirez	Angels	-9.12
Yu Darvish	Rangers	-9.11
Jerad Eickhoff	Phillies	-8.59

\*Minimum of 120 IP and 10 pitches per type to qualify

Table 11: Regression Results (ERA)

VARIABLES	(1) ERA	(2) ERA	(3) ERA	(4) ERA	(5) ERA
Age	0.899 [1.894]	-0.021 [1.888]	-0.127 [2.233]	2.603* [1.509]	0.313 [0.407]
Age Squared	-0.010 [0.033]	0.008 [0.032]	0.004 [0.038]	-0.039 [0.025]	
Fastball Velocity	-0.296 [0.209]			0.041 [0.458]	
Slider Velocity	0.174 [0.214]			0.200 [0.397]	
Curveball Velocity	-0.084 [0.191]			-0.149 [0.292]	
Change-up Velocity	0.205 [0.177]			-0.538** [0.211]	0.293** [0.141]
Fastball Horizontal		-0.092 [0.171]		0.675*** [0.251]	0.003 [0.147]
Slider Horizontal		0.113 [0.239]		-0.983*** [0.288]	0.038 [0.169]
Curveball Horizontal		-0.157 [0.229]		-0.049 [0.234]	
Fastball Vertical			0.082 [0.233]	0.921*** [0.266]	-0.244 [0.269]
Slider Vertical			-0.170 [0.158]	-0.345 [0.374]	
Curveball Vertical			0.160 [0.148]	-0.503* [0.258]	0.087 [0.163]
Change-up Vertical			0.198 [0.191]	-0.578* [0.325]	0.277 [0.210]
Fastball ICD				-0.358** [0.178]	-0.068 [0.089]
Curveball ICD				0.009 [0.037]	
Slider ICD				-0.123*** [0.030]	-0.017 [0.043]
Change-up ICD				0.023 [0.047]	
Constant	-11.425 [29.072]	-0.579 [28.150]	3.472 [32.447]	-5.850 [27.901]	-28.258 [17.135]
Observations	108	108	108	81	99
R-squared	0.1983	0.0447	0.0594	0.3733	0.1883
Root MSE	0.333	0.321	0.322	0.195	0.316
F-Stat	1.51	1.96*	2.32**	13.60***	1.90*

Robust standard errors in brackets

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

Table 12: Regression Results (FIP)

VARIABLES	(1) FIP	(2) FIP	(3) FIP	(4) FIP	(5) FIP
Age	0.586 [1.242]	1.364 [1.227]	0.474 [1.479]	4.190*** [0.879]	2.112 [1.874]
Age Squared	-0.006 [0.022]	-0.018 [0.021]	-0.002 [0.026]	-0.064*** [0.015]	-0.023 [0.030]
Fastball Velocity	-0.026 [0.176]			-0.033 [0.302]	
Slider Velocity	0.234 [0.149]			0.156 [0.119]	
Curveball Velocity	-0.147 [0.111]			0.066 [0.194]	
Change-up Velocity	-0.112 [0.088]			-0.457*** [0.146]	0.229 [0.159]
Fastball Horizontal		-0.072 [0.071]		-0.009 [0.173]	
Slider Horizontal		0.068 [0.119]		-0.639*** [0.172]	-0.055 [0.130]
Curveball Horizontal		-0.014 [0.145]		0.324*** [0.121]	0.047 [0.111]
Fastball Vertical			-0.301* [0.162]	0.334* [0.186]	-0.577** [0.272]
Slider Vertical			0.051 [0.091]	-0.200** [0.085]	0.014 [0.128]
Curveball Vertical			0.003 [0.053]	-0.245*** [0.091]	-0.062 [0.084]
Change-up Vertical			0.209* [0.117]	-0.296*** [0.107]	0.180 [0.152]
Fastball ICD				-0.235*** [0.036]	-0.122* [0.073]
Curveball ICD				0.018 [0.012]	
Slider ICD				-0.067*** [0.025]	-0.026 [0.029]
Change-up ICD				-0.058** [0.023]	-0.054* [0.030]
Constant	-3.974 [20.350]	-20.451 [18.417]	-6.096 [21.342]	-42.067** [18.338]	-52.276 [32.950]
Observations	108	119	108	81	93
R-squared	0.2212	0.0297	0.0362	0.4791	0.3027
Root MSE	0.223	0.241	0.222	0.0921	0.196
F Statistic	2.37**	2.28*	1.79	124.64***	1.74

Robust standard errors in brackets

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1



Table 13: Regression Results (Salary, Part 1)

VARIABLES	(1) Salary <sub>t+1</sub>	(2) Salary <sub>t+1</sub>	(3) Salary <sub>t+1</sub>	(4) Salary <sub>t+1</sub>
ICD				80,050.519 [312,049.813]
Age	12690496.472 [11207197.432]	6971175.774 [9966119.197]	5109815.165 [12380977.724]	14677662.716 [15616872.345]
Age Squared	-196,215.353 [197,477.681]	-98,120.451 [177,757.022]	-66,885.764 [221,102.846]	-227,662.139 [271,912.942]
Fastball Velocity	2119399.919 [1378614.055]			2445909.454 [2358431.664]
Slider Velocity	1853387.818* [938,276.311]			1745689.900 [1543658.208]
Curveball Velocity	-2318182.925*** [425,772.024]			-2580616.989*** [752,273.589]
Change-up Velocity	482,217.676 [945,912.813]			535,890.642 [1203794.598]
Fastball Horizontal		-381,432.900 [372,538.309]		162,557.403 [620,840.013]
Slider Horizontal		-1058414.673 [846,523.061]		-228,269.697 [1059170.015]
Curveball Horizontal		-3,911.456 [650,582.330]		39,993.569 [985,274.950]
Fastball Vertical			1329632.931 [2088125.092]	-700,118.138 [2221891.268]
Slider Vertical			576,011.652 [703,902.401]	269,574.543 [1127831.283]
Curveball Vertical			-505,489.559 [513,138.882]	214,866.800 [525,054.554]
Change-up Vertical			187,144.147 [1095526.163]	631,656.908 [890,590.342]
Constant	-4.050e+08** [1.874e+08]	-1.071e+08 [1.381e+08]	-9.747e+07 [1.798e+08]	-4.366e+08 [2.953e+08]
Observations	31	36	31	31
R-squared	0.604	0.243	0.280	0.637
Root MSE	6.873e+06	8.888e+06	9.269e+06	8.063e+06
F-Stat	8.26***	4.99***	2.85**	4.77***

Robust standard errors in brackets

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 14: Regression Results (Salary, Part 2)

VARIABLES	(1) Salary <sub>t+1</sub>	(2) Salary <sub>t+1</sub>
Age	5763917.624*** [1498719.523]	6527198.003*** [1433638.150]
Age Squared	-76,130.205*** [22,781.907]	-87,218.364*** [21,869.286]
Wins	291,661.133 [299,037.359]	
Innings Pitched	11,574.967 [43,743.699]	
Strikeout Rate	1277345.894* [649,373.168]	1604824.592** [611,987.832]
Walk Rate	-3332405.439** [1569619.150]	-2459205.351** [1029933.397]
Hit Rate	-1976093.401 [1689118.330]	
Home Run Rate	-8047489.133** [3398358.273]	-6760186.889** [3181263.054]
Earned Run Average	3867526.317 [2467777.660]	
Constant	-8.801e+07*** [29783598.213]	-1.035e+08*** [21918082.635]
Observations	71	71
R-squared	0.543	0.511
Root MSE	6.207e+06	6.217e+06
F-Stat	9.48***	14.87***

Robust standard errors in brackets

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1