Do Major League Baseball Pitchers Employ Optimal Mixed Strategies?

A Master's Thesis for the Department of Economics

Tufts University

by Jarrod Alexander Smith

May 2011

Abstract: The following paper analyzes pitch-level data to determine whether Major League Baseball pitchers use optimal mixed strategies. It builds upon an empirical strategy presented in the existing literature, while making novel adjustments to account for certain nuances of the pitcher-batter interaction. The results of this paper are in accordance with those of prior studies; namely, MLB pitchers do not engage in optimal play by throwing a disproportionately high number of fastballs. It concludes with the proposition that, to date, the pitcher-batter interaction has not been accurately modeled, and going forward studies should incorporate a pitch-sequencing valuation scheme in order to produce more credible results

Table of Contents

Section I: Introduction

Since its inception in the late 1930's, game theory has been regarded with increasing interdisciplinary scrutiny due to its ability to demystify the decision-making processes of individuals facing incentive-bearing dilemmas. In its essence the field consists of the study of [mathematically] well-defined interactions amongst agents for the purpose of characterizing optimal behavior. Despite its ability to provide both predictive and prescriptive insights into the behavioral responses of individuals confronted with complex multi-stage interactions, empirical evidence in support of the theory associated with even the simplest games has proven to be extremely elusive in practice.

Any test of the fundamental tenets of optimal play necessarily begins with a clear statement of the interaction of interest. This task includes the specification of: the players, the rules of the game, the actions available to each player, the payoffs under every potential game scenario, and the objectives. For simplicity researchers have historically confined empirical studies to simple two-player, simultaneous-move, zero-sum games with clearly defined (often binary) objectives. Regardless of the type of interaction being analyzed, the predictions characterizing equilibrium behavior rely upon the assumptions that all players are rational agents seeking to maximize their respective utilities, and possess a deep understanding of the rules, actions, and payoffs of the game. As such, any study seeking to test for optimal play, whether it be lab-based or field-based, will ideally analyze games whose participants satisfy these criteria. Although they differ in the particulars, a variety of barriers exist which complicate the detection of optimal play in both lab-based and field-based analyses of strategic interaction.

Evidence of Optimal Play is Elusive in the Lab and Field

Although carefully designed experiments offer researchers the advantage of structuring tractable games with a limited number of players, actions, and payoffs, they have generally been unable to conclude that participants engage in optimal play as dictated by theory. Oftentimes the volunteers for these studies are unfamiliar with the games and must learn them during the course of the experiment. As novices they may be unable to effectively optimize their choice of actions due to confusion or tentative gameplay. Further confounding the results is the trivial nature of the payoffs provided to laboratory participants; even if the players have the capacity to identify and play optimal strategies, the incentives provided are often insufficient to prompt them to do so. These barriers can be overcome by analyzing real-world, high-stakes interactions in which the agents are seasoned professionals in their respective fields.

Although they are able to circumvent the difficulties associated with inexperienced agents typical to lab experiments, studies analyzing field data have similarly yielded inconclusive results in verifying the implications of the theory. The primary drawbacks to using field data stem from the difficulty of observing and recording distinct actions and payoffs. Unlike in the laboratory, naturally-arising interactions (or "games") with nontrivial consequences often consist of actions that vary across multiple dimensions, and payoffs contingent upon a variety of factors (many of which may be unobservable). Moreover, the interactions being studied often lack a well-defined set of rules to govern the players, which further complicates the task of identifying optimal play. In response to the pervasive difficulties of identifying optimal play, researchers have shifted their focus to less-traditional venues in hopes of mitigating the complications associated with standard lab and field-based studies.

Seeking Optimal Strategies in Professional Sports

One promising trend that has recently garnered attention is the behavioral analysis of professional athletes. The arena of professional sport possesses a variety of qualities which make it an attractive venue for testing the implications of game theory. Professional sports are governed by specific rules and objectives known to the players. Due to fierce competition at amateur levels, professionals must possess expert-level knowledge of the rules, actions, payoffs, and objectives of the sport. Over the span of a season athletes are repeatedly confronted with non-trivial choices that will influence their future salaries by millions of dollars. For the aforementioned reasons, professional athletes satisfy the typical assumptions of agents upon which predictions of optimal play are grounded.

Large supplies of detailed and freely-available data further facilitate the game theoretic analysis of interactions within professional sports contests. Many professional sports leagues, teams, and private third-party companies collect and disseminate play-by-play accounts of individual games free of charge, making it possible for researchers to break down complete games into sequences of distinct actions. The only remaining difficulty is to identify sub-interactions with clearly defined players, actions, and outcomes. The existing literature focuses primarily on sports featuring isolated one-on-one interactions with binary measures of success; interactions such as these readily conform to standard theoretical two-player zero-sum games and serve as a natural starting point for empirical analysis .

Studies analyzing professional sports data typically test whether athletes maximize expected payoffs according to two clear and testable predictions of the minimax theorem: 1) actions are chosen in proportions such that the expected payoffs are equal across all actions, and

2) agents exhibit zero serial correlation in their action choices. Walker & Wooders consider the case of serve-and-return play in tennis matches at Wimbledon, and conclude that although players do choose strategies such that the payoffs are equal across actions, they do not fully randomize with respect to actions, and thus certain actions exhibit serial correlation (2001). Palacios-Huerta finds evidence in penalty-kick data from European professional soccer leagues that suggests soccer players (both goalkeeper and kicker) choose strategies which satisfy both criteria of the minimax theorem (2003).[1] Additionally, Chiappori et al. explore the implications of player heterogeneity in soccer penalty-kicks, and find evidence in European professional soccer league data that players use optimal mixed-strategies (2002).

Exhaustive analysis of soccer and tennis data has prompted researchers to test for optimal play in more complex sports sub-games, with baseball's pitcher-batter interaction drawing interest. Due to the variety and complexity of the actions, outcomes, and payoffs associated with a plate appearance, testing whether Major League Baseball (MLB) pitchers optimize with mixed-strategies poses modeling difficulties not present in the analysis of penalty-kicks and tennis serve-and-return play. In the two studies that have investigated this issue to date, both Kovash & Levitt (2009) and Weinstein-Gould (2009) respectively find evidence which suggests that MLB pitchers fail to adhere to the prescriptions of the minimax theorem.

The contents of this paper are organized as follows. Section II provides a critical review of the existing studies which test for the use of optimal mixed-strategies by MLB pitchers. Section III sets forth a simple theoretical model of the pitcher-batter interaction and characterizes optimal play. Section IV introduces an empirical strategy for testing for optimal play and highlights the distinction between the strategy used in this paper and those used in preceding

[1] It is worth noting that this is the first study which successfully indentifies serial independence of actions in a field setting.

studies. Section V describes the source and structure of the data, which contains key variables unique to the dataset. Section VI presents the results and their implications, and Section VII concludes with a discussion of issues which must be addressed in future research concerning the pitcher-batter interaction.

Section II: Literature Review

Comprehensive Overview of Studies Analyzing Pitcher Strategy

As stated in the introduction, only two studies to date have attempted to test for optimal mixed-strategies amongst MLB pitchers. The most notable game theoretic investigation of the pitcher-batter interaction was conducted in 2009 by Kenneth Kovash and Steven Levitt and is detailed in their paper "Professionals Do Not Play Minimax." In this study the authors develop an empirical framework for the analysis of pitch-level data and conclude that MLB pitchers fail to adhere to optimal mixed-strategies as prescribed by the minimax theorem. The dataset used was purchased from Baseball Info Solutions, a private company that specializes in collecting and analyzing baseball data for their clients, which include professional teams, memorabilia companies, agents, and academic researchers.

The dataset contains all pitches (roughly 3.5 million) thrown over the 2002-2006 seasons, and features variables identifying the pitch type, pitch result, number of outs, inning, count[2], runners on base, and identity of both the pitcher and batter. In modeling the pitcher-batter interaction as a two-player, simultaneous-move, zero-sum game the authors designate pitch type

---

[2] The count indicates the number of balls and strikes already thrown in a given plate appearance. For example, a 1-2 count indicates the pitcher has thrown one ball and two strikes prior to the current pitch.

as the sole choice (action) variable. Isolating pitch type as the only choice variable is clearly a simplification as pitchers also actively choose where to locate their pitches, but it is a seemingly necessary one since including location as a choice variable presents difficulties arising from the fact that even the most effective pitchers lack the command to deliver every pitch to the exact intended location. Additionally, the dataset used in this study does not contain variables identifying pitch location, so allowing for pitchers to choose locations would be impossible.

Although MLB pitchers throw a variety of different pitches, the authors collapse all pitches into four mutually exclusive categories: fastball, curveball, slider, and changeup. All pitches initially coded as forkball, knuckleball, pitchout, screwball, sinker, and unknown, which cumulatively account for 6% of the total pitches thrown, are dropped from the dataset. Of the remaining pitches, 64.33% are fastballs, 9.53% are curveballs, 13.62% are sliders, and 12.52% are changeups. Due to the relative importance of pitch type to the analysis, the authors cross-check the coding of pitch types with another dataset organized by STATS Inc., and find that the coding matches on over 90% of observations, with the majority of discrepancies occurring on off-speed[3] pitches.

Kovash & Levitt's primary focus is testing whether pitchers choose pitch types such that the average outcome from throwing each pitch type is equalized, which requires the designation of some measurement that captures the outcome of any given pitch. To this end the authors enlist the use of OPS, a commonly reported statistic that measures offensive productivity by taking the sum of a batter's on-base percentage (OBP) and slugging percentage (SLG)[4]. Due to the fact that OPS is the sum of two distinct offensive averages (with different denominators) it has no simple

---

[3] Any pitch that is not a fastball is commonly referred to as an off-speed pitch.
[4] OBP measures the frequency with which a batter reaches base (as a percentage of plate appearances), while SLG measures how many total bases per at-bat (base-on-balls and hit-by-pitch are not included in the numerator or denominator of this calculation) the batter achieves.

intuitive meaning, though it is clear that a higher OPS is indicative of increased offensive production. Since the pitcher-batter interaction is zero-sum, the effectiveness of any given pitch is related inversely to OPS, with the most effective pitch producing an OPS of zero.

The calculation of OPS for a single pitch requires that pitch to generate a specific result (ie. out, base-on-balls, single, double, triple, home run, etc.), so Kovash & Levitt necessarily drop all non-terminal pitches (pitches that do not end the plate appearance) in order to assign outcomes to each observation included in the analysis. As noted by the authors, one alternative to dropping non-terminal pitches is to code them according to the final result of the plate appearance, although this strategy obfuscates the relationship between pitch efficacy and OPS. For example, consider a five-pitch at-bat which begins with a fastball thrown for a strike, and concludes with the batter hitting a triple on a particularly poor curveball. In this case the first-pitch fastball will be assigned the relatively high OPS of 4.000, which is likely a misrepresentation of its value as a pitch.

After trimming the dataset and defining all necessary variables the authors develop an empirical strategy to test the hypothesis that MLB pitchers optimize across pitch types. They begin with a simple model and build upon it in piecewise fashion in order to address the sensitivity of their estimates to the inclusion certain controls. The first specification, which is a simple regression of OPS on indicator variables for each pitch type (with changeup as the base group) generates the following coefficients: .094 for fastball, -.060 for slider, and -.064 for curveball, all of which are statistically significant at the 1% level. The authors interpret the large and statistically significant positive coefficient on the fastball indicator as evidence that pitchers throw too many fastballs, with the average terminal fastball yielding an OPS that is 94 points[5]

---

[5] When discussing offensive statistics baseball statisticians generally refer to 1/10 of 1% as a “point”

higher than the average terminal changeup. The converse goes for sliders and curveballs, which initially appear to be the most effective pitches in a pitcher's toolbox.

The authors then present the results of four additional regressions in which they phase in a variety of controls in cumulative stages. The second specification adds count fixed-effects; the third specification adds fixed-effects for innings, outs, and number of runners on; the fourth specification adds pitcher, batter, and pitcher*batter interaction fixed-effects; and the fifth and most comprehensive specification adds fixed-effects for the pitcher*batter*count interaction. The estimates on the curveball and slider indicators decrease in magnitude with the addition of controls and the direction changes from negative to positive, although neither is statistically significant in the most saturated specification. The estimate on the fastball indicator retains its positive sign and statistical significance at the 1% level throughout all iterations, though the magnitude ranges from .041 to .073, with .073 occurring in the most saturated specification.

Due to the consistently positive and statistically significant effect of fastball on OPS across all specifications the authors conclude that MLB pitchers rely too heavily upon fastballs, and would do better to throw fewer since they yield the worst outcome on average. The varying estimates of the impact of off-speed pitches across different functional forms preclude them from arriving at any strong conclusions concerning whether curveballs and sliders are more or less effective than changeups on average. Since an OPS differential across pitch types is a relatively abstract measurement and does not directly address the degree to which overreliance on the fastball hurts pitchers, the authors seek to provide some context in order to make the implications of their results more accessible. By citing a formula developed by sabermetrician[6] Dan Fox which suggests than a incremental increase of a single OPS point generates roughly 1.8

[6] Sabermetrics is a movement amongst baseball statisticians which is concerned with the creation of more accurate and precise performance measures than the traditional statistics. SABR is an acronym for Society of American Baseball Researchers.

additional runs over the span of a season, the authors conclude that a pitching staff could reduce the numbers of runs allowed by approximately 15 runs per season by throwing 10% fewer fastballs, although they do not specify what pitches (and in what proportions) should be thrown instead of the fastballs.

The only other existing study that tests whether MLB pitchers adopt optimal mixed strategies is detailed in Weinstein-Gould's 2009 paper "Keeping the Hitter Off Balance: Mixed Strategies in Baseball." This study analyzes pitch-level data from the 2002 MLB season purchased from sports data collection company Tendu, which sells game data and analytical software to MLB teams, most notably the New York Mets and Oakland Athletics (Byous). Although the dataset contains variables for pitcher and hitter identity, pitch type, pitch result, and at-bat result, it lacks variables describing the game situation at the time of the pitch, such as outs, pitch count, and number of runners on base.

In addition to lacking situational variables, the dataset was coded manually by former college and professional players who viewed footage of television broadcasts, and thus it is likely to contain errors. It is worth noting that the proportion of fastballs to total pitches is nearly identical to that calculated in Kovash & Levitt's dataset, although the proportions for off-speed pitches differ by 1-2% each. This is to be expected, as distinguishing between a fastball and an off-speed pitch is relatively straightforward, while distinguishing between a curveball, slider, and changeup requires the ability to discern more subtle differences across pitches.

Much like Kovash & Levitt, Weinstein-Gould models the interaction between the pitcher and batter as a two-player, zero-sum, simultaneous-move game with pitch type as the sole choice variable, but this is where the similarity between the two models ends. Rather than designating OPS as the measure of pitch success, the author defines three distinct outcome variables. The first outcome variable, *Pitch*, is a binary measure of the immediate success of the pitch, which takes a value of 1 if the pitch results in a strike or an out and 0 if the pitch results in a ball, hit, or walk. The second outcome variable, *OnBase*, looks at the success of the plate appearance rather than the immediate success of the pitch, and takes a value of 1 if the batter reaches base and 0 if the batter does not reach base. The third outcome variable, *wOnBase*, is similar to the second in that it also measures the success of the plate appearance, but it weights the result based on its likelihood to produce runs. For example, while *OnBase* takes on a value of 1 for both a single and a triple alike, *wOnBase* assigns a higher value for a triple since it is substantially more likely to produce a run.

Weinstein-Gould drops a number of pitches from his dataset, so that he only includes the first pitch from each plate appearance in his regressions. He reasons that by doing so, all pitches will occur at the same count (no balls and no strikes), and will not be contingent upon the pitch type thrown immediately prior. This adjustment leaves the dataset with 79,107 observations, spanning 135 pitchers and 855 batters. Before testing for optimal mixed-strategies the author proposes that, due to the heterogeneity of batters, pitchers will alter strategies across hitters. To test this claim he regresses an indicator variable for each pitch type on fixed-effects for both hitters and pitchers. These regressions yield highly significant F-statistics for joint significance of the hitter fixed-effects for all pitch types excluding split-finger fastballs and knuckleballs, supporting his hypothesis that a given pitcher's strategy is uniquely tailored to the opposing batsman.

After verifying the batter-dependant nature of pitchers' strategies Weinstein-Gould presents his empirical framework for testing the null hypothesis that MLB pitchers select pitch types optimally. His method consists of regressions for each of the outcome variables on pitcher*batter interactions and pitcher*batter*pitch interactions, which, unlike the specifications used by Kovash & Levitt, allows the payoff of each pitch type to vary across each pitcher-batter matchup. All observations in which the pitcher and batter did not meet at least 10 times are dropped in order to impose a minimum cell size on the pitcher*batter interactions.

Unsurprisingly, the results are dependent upon which outcome variable is used. When *Pitch* is designated as the outcome variable, the pitcher*batter*count interaction terms are jointly insignificant with a p-value of .344, suggesting that pitchers select pitch types optimally with respect to the immediate outcome of the first pitch. On the other hand, the tests for joint significance of the pitcher*batter*count interactions yield p-values of .031 and .095 when *OnBase* and *wOnBase* are designated as the outcome variables, suggesting that pitchers do not select pitches optimally with respect to the final outcome of the plate appearance. The regressions which use *OnBase* and *wOnBase* as the dependent variables are more akin to the specifications used in the Kovash & Levitt study since they assign pitch values according to the outcome of the plate appearance, and therefore the results of the respective papers are in accordance. Since Weinstein-Gould allows for the return to each pitcher to vary across every possible pitcher-batter combination, he is unable to conclude which pitches are being thrown too often or too sparingly on average.

Critical Analysis

Although Kovash & Levitt succeed in establishing a framework for testing whether MLB pitchers optimize across pitch types, certain details of their model leave the validity of its conclusions open for debate. One aspect of the empirical strategy that drew a great deal of objection amongst the sabermetric community is the use of OPS as the outcome variable (Tango 2009). The contention arises due to the fact that OPS notoriously undervalues the contribution of a base-on-balls (BB). For example, a pitch that results in a single is assigned an OPS of 2.000, while a pitch that results in a BB is assigned an OPS of 1.000. Although a single is, on average, more valuable than a BB (i.e. a single with a runner on second will often produce an immediate run while a BB with a runner on second fails to advance the runner) it is certainly not twice as productive.

Additionally, Kovash & Levitt miscalculate OPS to further undermine the value of a BB by not accounting for the fact that OPS is actually the sum of two averages, OBP and SLG, with distinct denominators. Their calculation implicitly codes a BB with an SLG of 0.0000, when in fact a plate appearance that results in a BB is not included in the denominator of SLG, so it will not penalize SLG, but rather leave it unchanged. Since off-speed pitches like curveballs, sliders, and changeups are more difficult to command than fastballs (and perhaps more likely to result in a BB), it may be the case that the use of OPS as the outcome variable mutes the downside to off-speed pitch types even when it is calculated correctly. The miscalculation, which serves to further reduce the value of a BB, will only exacerbate this bias.

An additional drawback to the choice of OPS as the outcome variable relates to its esoteric nature as a statistic. In modeling any two-person, zero-sum interaction one must designate a specific outcome which both players attempt to manipulate through the selection of distinct actions available to them. Although OPS is certainly a valid proxy for both pitcher and batter success, the proposition that MLB pitchers (and batters) explicitly try to minimize (and maximize) the OPS of the batter seems unrealistic. It is more likely that the ultimate objective of each player is to win the game as a whole rather than each individual at-bat, which is accomplished from the pitcher's perspective by run prevention and from the batter's perspective by run production. In this case a more appropriate outcome variable would be one that weights each outcome according to its ability to produce runs, much like *wOnBase* from Weinstein-Gould's third specification.

Another aspect of Kovash & Levitt's empirical strategy that leaves its results susceptible to criticism is the fact that only pitches which end the plate appearance are analyzed. This strategy neglects the possibility that a specific pitch type may perform especially well *when it is not a terminal pitch*. For example, suppose a fastball frequently results in a strike when it does not end the plate appearance, but performs comparatively worse than off-speed pitches when it does end the plate appearance. In this case an analysis of only terminal pitches will mask the value of the fastball's ability to record strikes which do not terminate the plate appearance, thus making it appear less effective than breaking pitches. It is worth noting that Kovash & Levitt do provide tables which indicate that the average OPS of a fastball is not substantially different from the average OPS of breaking pitches when the pitches do not end the plate appearance, but as previously mentioned, the ability of OPS to capture the value of a non-terminal pitch is dubious.

The empirical strategy used in the Weinstein-Gould paper also contains aspects which threaten the validity of the results. The most glaring issue is the lack of controls included in the regressions, which is entirely a function of the dataset. Since his dataset lacks information describing the game situation, Weinstein-Gould is unable to condition on situational variables like count, inning, outs, and runners on. The omission of relevant controls creates the potential for bias in the estimates of interest, which are the coefficients on the pitcher*batter*pitch interaction terms. For example, if pitchers tend to throw more fastballs with runners in scoring position and hitters perform better with runners in scoring position, then the estimate on fastball may contain an upward bias, making it look as if fastballs are less effective than they truly are.

Perhaps the most important determinant of pitch type that is not controlled for is count; conventional baseball wisdom indicates that pitchers throw pitches with different proportions according to whether they are faced with an advantageous or disadvantageous count. The author deals with this problem by choosing to only analyze the first pitch of each plate appearance, thus guaranteeing that all pitches analyzed are thrown at the same count: no balls and no strikes. While this strategy eliminates any potential bias due to omitting count fixed-effects, it spawns new complications concerning the link between the action choice and outcome. Consider the specifications in which *OnBase* and *wOnBase*, which are both defined according to the final result of the plate appearance, are designated as the outcome variable. In a one-pitch at-bat the link between the first pitch and the outcome is clear, but for longer at-bats the influence of the first pitch on the final result diminishes with each subsequent pitch. In a seven-pitch at-bat resulting in a home run it is likely that the first pitch plays a minor (or perhaps nonexistent) role in allowing the home run, thus testing whether pitchers select the first pitch optimally with respect to the outcome of the plate appearance makes little sense and may contribute to findings which contradict the theory.

Section III: Theoretical Model

In accordance with the existing literature this paper presents the pitcher-batter interaction as a two-player, simultaneous-move, zero-sum game in which the pitcher chooses which pitch to throw and the batter chooses which pitch to anticipate. The game is considered simultaneous-move because hitters do not have enough time to choose their action after the pitch is thrown; when facing an 87 mph pitch (which is the average pitch speed over the 2009 and 2010 seasons) the hitter has roughly .47 seconds to respond, thus from a practical standpoint the batter does not observe the pitcher's action prior to selecting his own. Additionally the game is assumed to be played with complete information since all professional teams retain extensive scouting departments to track the historic tendencies and performances of both pitchers and hitters (and even umpires), so detailed accounts of the past strategies and payoffs of the opponent are known to all players (McCauley et al.).

The following example motivates why pitchers must throw pitches in a proportion that equalizes outcomes across pitch types in order to optimize their expected payoff. Consider a simplified version of pitcher-batter interaction in which the pitcher chooses to throw either a fastball or an off-speed pitch and the batter chooses to anticipate either a fastball of an off-speed pitch. The following payoff matrix characterizes the associated payoffs with each pairing of actions:

		Batter	
		Fastball	*Off-speed*
Pitcher	*Fastball*	φ_{ff}	φ_{of}
	Off-speed	φ_{fo}	φ_{oo}

Although the outcome measure of the actual pitcher-batter interaction is a complex and nuanced concept, the preceding example assumes that the batter simply attempts to reach base, while the pitcher attempts to record an out (similar to Weinstein-Gould's second specification which uses *OnBase* as the dependent variable), so φ_{ij} represents the probability that the batter reaches base when he anticipates pitch type i and the pitcher selects pitch type j. We can safely assume that the batter is able to hit more effectively when he correctly anticipates the pitch type, so that $\varphi_{ff} > \varphi_{fo}$, $\varphi_{ff} > \varphi_{of}$, $\varphi_{oo} > \varphi_{of}$, and $\varphi_{oo} > \varphi_{fo}$, thus a pure strategy equilibrium does not exist.

In this case both pitcher and batter must adhere to strategies in which they randomize across both action choices with certain probabilities. In order to maximize his expected payoff, each player must choose to mix actions so that the opponent is indifferent between his available actions. Let ρ_f represent the probability that the pitcher throws a fastball. The pitcher must choose ρ_f to satisfy the equality $\rho_f \varphi_{ff} + (1- \rho_f)\varphi_{fo} = \rho_f \varphi_{of} + (1- \rho_f)\varphi_{oo}$, where the left side of the equation represents the batter's probability of reaching base if he chooses to anticipate a fastball and the right side represents the batter's probability of reaching base if he chooses to anticipate an off-speed pitch. If the equality does not hold then the batter would do better to choose a pure strategy in favor of the action that yields the higher probability of reaching base. The same logic applies to the batter, who must choose β_f, the probability that he anticipates a fastball, in order to satisfy the equality $\beta_f\varphi_{ff} + (1- \beta_f)\varphi_{of} = \beta_f\varphi_{fo} + (1- \beta_f)\varphi_{oo}$, so that the pitcher must be indifferent between throwing a fastball or throwing an off-speed pitch.

In this case, if each player is maximizing his expected payoff, then both pitcher and batter must be indifferent between actions since the payoffs across actions must be equalized. If one action choice yields a higher expected payoff than the other, then the player can enhance his return by choosing that action more frequently until the disparity in payoffs between the two actions is eliminated. If the disparity is never eliminated such that one action always yields a higher expected payoff than the other, then the action with the lower expected payoff will never be played. The end result of this process is that in equilibrium, all actions that are played with a positive probability must yield the same outcome on average. Although the preceding example is admittedly a simplification, the fundamental result of equalized payoffs across actions can be extended to the actual pitcher-batter interaction, which is characterized by multiple actions choices (pitch types) and a more complex, non-binary outcome measure. The next section details the dataset used in this study, while Section V presents an alternative empirical strategy for testing whether expected outcomes are equalized across action choices using MLB pitch-by-pitch data.

Section IV: Data

Source

The data analyzed in this study were generated by MLB with a technology known as Pitch f/x. Pitch f/x is a system developed by sports technology company Sportsvision which captures a wide variety of physical measurements and situational information concerning pitches thrown in MLB games. MLB rolled out the use Pitch f/x on a limited basis in the 2007 season, and had fully implemented the system in all 30 stadiums by the beginning of the 2008 season (Nathan 2007). To date, no study investigating the use of optimal mixed-strategy amongst MLB pitchers has analyzed Pitch f/x data.

The primary purpose for the integration of Pitch f/x technology into MLB stadiums was to provide fans without access to televised broadcasts a means of following games on a pitch-by-pitch basis in real-time. The data collected by Pitch f/x are immediately posted to MLB.com's "Enhanced Gameday" website, which features a graphical interface illustrating the locations and result of each pitch, enabling fans to "watch" the game without actually tuning in to a televised broadcast (Newman). The data are also stored permanently to an archive section of MLB's website with distinct inning files for each game, and is available free of charge.

In the simplest sense Pitch f/x is a system of three high-speed cameras (30 frames-per-second), all of which are strategically located in different areas of the stadium, just behind the outfield fence. The cameras are calibrated by technicians prior to every game in order to triangulate the strike zone, which is a somewhat nebulous concept defined in vague terms by MLB (Newman). Although the horizontal range of the strike zone is explicitly defined as the width of home plate, which is exactly 17 inches across, the vertical range is more open to interpretation and is laid out in the rulebook as follows: "the upper limit… is a horizontal line at the midpoint between the top of the shoulders and the bottom of the uniform pants, and the lower level is a line at the hollow beneath the knee caps" ("Official Rules"). In order to deal with this potentially confusing definition, system technicians assign unique vertical coordinates for the top and the bottom of the strike zone for each player during batting practice prior to the start of play. After a pitch is captured by the cameras the footage is immediately transmitted to a processing truck stationed at the loading dock of the ballpark. The processing truck contains three computers which calculate a variety of measurements associated with each pitch, and then post this information online to both the "Enhanced Gameday" feature and the archival data section of MLB.com (Newman).

Although the information collected by the Pitch f/x system is conveniently located online free of charge, collecting data for a full season is an extremely cumbersome process due to the fact that each inning has its own distinct file (a full season is roughly 24,570 innings, excluding extra innings). In response to the large up-front costs of assembling the data, a number of prominent baseball researchers with experience in computer programming have created scripts that collect the data from its original source and organize it into workable databases, some of which are made freely available online. While these scripts necessarily vary in their respective details, the general process is always similar: they use a simple programming language to create a program which pulls the data from the website and stores it in text files, which are then parsed into single or multiple SQL databases.

The dataset used in this paper was obtained in the form of two seasonal SQL databases from a website run by Joe Leftkowitz, an amateur baseball researcher.[7] Each SQL file contains every Pitch f/x observation from the 2009 and 2010 seasons, excluding postseason play. The databases respectively contain 717,254, and 710,329 observations. Due to the gargantuan size of the datasets and the intricacies of the Pitch f/x system, it is unsurprising that the raw datasets contains errors which must be removed prior to analysis.

Errors in the Dataset

The most common error present in the dataset arises as a result of Pitch f/x system malfunctions, which are inevitable given the intricacies of the system. The databases for both of the aforementioned seasons contain a number of observations in which certain pitch-level

[7] These datasets, and more, can be found at Joe Leftkowitz's website, located at: http://www.joelefkowitz.com/index.php

variables that can only be measured by high-speed video footage, such as start speed, end speed, vertical location coordinate, horizontal location coordinate, and others, are recorded as null. For these same observations other pitch-level variables which can be recorded without the system of cameras, such as pitch type, pitch designation as a strike or a ball, the count, and others, are accurately recorded. It seems likely, due to the fact that the only measurements not properly recorded are those which require high-speed video footage to calculate, that these errors are a result of a camera malfunction.

Another type of error present in the dataset again relates to only those pitch-level variables which cannot be read in without the system of cameras. Each observation contains a date and timestamp which records the date, hour, minute, and second at which the pitch was recorded by the system. There are a number of pitches thrown within a given game that posses the same timestamp. These pitches often occur during different at-bats, and in addition to sharing the timestamp, they also share all variables which can only be read in by the system of cameras. Variables such as count, pitch type, etc. are presumably accurately recorded for these pitches.

Since many of the variables that are rendered null or inaccurate by the aforementioned errors are instrumental to my analysis, these observations have been deleted from the dataset. Additionally, I have deleted every observation from half-innings in which at least one of these errors is present, so that the seasonal samples consist of only pitches from half-innings in which the Pitch f/x system accurately recorded every single pitch. This extra precaution addresses the possibility that even the pitches for which all variables were recorded during a half-inning containing at least one system malfunction may contain inaccuracies due to technological complications. After deleting all such observations the seasonal datasets are left with 665,294 and 665,759 observations for the 2009 and 2010 seasons respectively.

Variables of Interest

In addition to containing error pitches, the datasets also contain a large number of variables which are not explored in this study. The Pitch f/x cameras are able to track the path of any given pitch from the release point of the pitcher's hand to its final destination, whether that is the glove of the catcher or the bat of the hitter. The high frame-rate of the footage allows the for the calculation and estimation of a large number of measurements related to each pitch which were previously unobservable, such as location, spin direction, spin rate, break angle, and others. The system records roughly 41 variables for each observation, but in order to reduce the size of the datasets I have dropped all variables not used in this study, which leaves me with a total 29 variables for each pitch.

The following table presents a description of all relevant variables contained in the raw SQL files. All variables used in this study that are not explained in the following table are generated from information captured by these variables, and will be carefully explained in the empirical strategy and results sections. Level describes the stratification along which these variables remain constant. To clarify, game-level variables do not change within games; inning-level variables do not change within innings; atbat-level variables do not change within atbats; and pitch-level variables may potentially change across every pitch.

Variable Name	Type	Level	Description
gid	string	game	identifies date, home team, and away team
stadium	numeric	game	unique id for stadium
ump	numeric	game	unique 6-digit id for home-plate umpire
pitching_team	string	inning	identifies team in field
batting_team	string	inning	identifies team at-bat
inning	numeric	inning	identifies inning
pitcher	numeric	atbat	unique 6-digit id for pitcher
p_throws	string	atbat	"L" if lefty, "R" if righty
batter	numeric	atbat	unique 6-digit id for batter
batter_handedness	string	atbat	"L" if lefty, "R" if righty
sz_top	numeric	atbat	coordinate of top of strike zone for given batter
sz_bot	numeric	atbat	coordinate of bottom of strike zone for given batter
atbat_num	numeric	atbat	number of atbat for given pitcher
atbat_result	string	atbat	description of result of given atbat
sv_id	numeric	pitch	date and timestamp for given pitch
pitch_count	numeric	pitch	pitch number for given pitcher
atbat_pitch_num	numeric	pitch	pitch number for within a given atbat
pitch_type	string	pitch	identifies pitch type
pitch_result	string	pitch	description immediate result of given pitch
balls	numeric	pitch	number of balls thrown in atbat prior to pitch
strikes	numeric	pitch	number of strikes thrown in atbat prior to pitch
outs	numeric	pitch	number of outs recorded prior to pitch
on_first	numeric	pitch	1 if runner on first, 0 if not
on_second	numeric	pitch	1 if runner on second, 0 if not
on_third	numeric	pitch	1 if runner on third, 0 if not
start_speed	numeric	pitch	speed (in mph) of pitch as it leaves pitcher's hand
end_speed	numeric	pitch	speed (in mph) of pitch as it crosses plate
px	numeric	pitch	horizontal coordinate of pitch as it crosses plate
pz	numeric	pitch	vertical coordinate of pitch as it crosses plate

Section V: Empirical Strategy

Choice Variable

In accordance with the modeling strategies set forth by both Kovash & Levitt and Weinstein-Gould, I assume that pitchers' strategies consist of selecting only pitch type. A richer model would include pitch location as a choice variable, but this presents complications stemming from the fact that pitchers often fail to deliver the pitch precisely to the intended location. This could potentially lead to a false rejection of optimal play due to faulty execution of pitch strategy rather than sub-optimal selection of actions. This problem does not arise when pitch type is regarded as the sole choice variable; a pitcher cannot unintentionally throw a curveball when he intends to throw a fastball. The coefficients on the indicator variables for pitch type are of primary interest in this study; if pitchers truly optimize across actions then each pitch type should have no impact on the outcome measure, relative to the base group.

A pitcher's strategy is comprised of the following four actions: fastball, curveball, slider, and changeup. The raw Pitch f/x data provided by MLB codes each pitch as one of 18 different type pitches. Depending on the initial pitch type code I assign each pitch into one of the four mutually exclusive 'umbrella' groups or drop it from the dataset. For consistency I group pitches according to the same scheme employed by Kovash & Levitt. Fastball includes pitches initially coded by the Pitch f/x system as fastball, two-seam fastball, four-seam fastball, and cut fastball; curveball includes curveball and knuckle-curve; slider includes only slider; and changeup includes changeup and split-finger fastball. The remaining pitch types, which cannot be neatly collapsed into one of the four 'umbrella' groups, are classified as other, and dropped from the dataset prior to analysis. The following table identifies each pitch type's proportion of total pitches thrown in the dataset.

Pitch Type	**2009**		**2010**	
	Number of Pitches	Frequency	Number of Pitches	Frequency
Fastball	395704	59.48%	364390	54.73%
Slider	111901	16.82%	95789	14.39%
Curveball	61744	9.28%	61693	9.27%
Changeup	79035	11.88%	86059	12.93%
Other	16910	2.54%	57828	8.69%

The table highlights the fact that the proportion of pitches designated as other increased drastically from the 2009 to the 2010 season, which suggests structural changes in Pitch f/x pitch type coding over time. The following table breaks down all pitches designated as other in order to explore the trend.

Pitch Type	**2009**		**2010**	
	Number of Pitches	Frequency	Number of Pitches	Frequency
Sinker	10973	1.6493%	48603	7.3004%
Knuckleball	2386	0.3586%	4153	0.6238%
Knuckle-Curve	0	0.0000%	994	0.1493%
Screwball	0	0.0000%	109	0.0164%
Forkball	0	0.0000%	217	0.0326%
Eephus	0	0.0000%	89	0.0134%
Intentional Ball	3019	0.4538%	3137	0.4712%
Pitchout	441	0.0663%	517	0.0777%
Balk	1	0.0002%	1	0.0002%
Unknown	90	0.0135%	8	0.0012%
Total	**16910**	**2.5417%**	**57828**	**8.6860%**

The preceding table indicates an increasing sophistication of the Pitch f/x system to distinguish between pitches that are not fastballs, sliders, curveballs, or changeups. The most striking change is the roughly 340% increase in the number of pitches initially coded as sinkers over the 2009 to 2010 seasons. The falling proportion of pitches designated as fastballs and sliders over the two seasons suggests that the majority of these sinkers are coming out of the fastball and slider categories. This proposition is bolstered by the fact that sinkers are roughly the same speed as fastballs and sliders, and often mimic the downward movement that characterizes both pitches. The increasing likelihood that a sinker is properly identified may potentially have ramifications on the estimates of fastball and slider since the pitches that comprise these groups have changed over the two seasons. Additionally, the system codes four new pitch types that were not present during the 2009 season, though this change is unlikely to affect the estimates of interest since the new pitch types only account for approximately .21% of the total pitches thrown during the 2010 season.

Outcome Variable

The primary deviation of the empirical strategy employed in this paper from those suggested in the existing literature lies in the designation of Run Expectancy (RE) as the outcome variable. RE is a measure of offensive production developed in the 1970s by revered sabermetrician Pete Palmer and is an essential component in the calculation of Wins Above Replacement (WAR), a statistic used to evaluate players by a number of MLB general managers (DiFucci). Although it is known by a variety of monikers, the measure was initially introduced under the name linear weights due to the fact that it assigns weights to certain events according to the event's ability to generate a run. For example, the generally accepted RE for a single (unconditional on the situation) is roughly .47, meaning that the average single creates .47 runs for the offensive team prior to the end of the inning. Accurate calculation of the RE of any given event requires vast play-by-play databases which depict every play from every game over a specified period of time.

An RE score for a given event is calculated as follows. For simplicity let us assume the event is a single. First, the base-out state (number of outs and location(s) of runner(s) on base) prior to the occurrence of the single is determined, and the average number of runs scored prior to the end of the inning (RBOI) from that specific base-out state is calculated. Next, the base-out state following the single is determined, and the RBOI from the resulting base-out state is calculated. The difference between the RBOI of the beginning and resulting states plus the number of runs scored on the single is the 'runs added' for *that specific single*. Unconditional RE for a single is then calculated by averaging the 'runs added' for all singles occurring in a specified time frame (Birnbaum). Since RE measures are generated by empirically-based calculations they are only able to capture the run production of certain events for give time period, although the values tend to remain relatively stable over season-long intervals (Klaassen).

RE possesses a comparative advantage over other measures of production for a variety of reasons. While it is impossible to know the utility functions of major league pitchers and batters (or whether they are homogenous for that matter), it is likely that all players attempt to maximize the probability of winning to some degree. Most players are eligible for annual bonuses tied to both personal and team performance, and players judged to be most instrumental in wins are often rewarded with generous long-term contracts. The most conspicuous means of directly contributing to a win is through run prevention and production for pitchers and batters respectively, thus it is appropriate to assume that within the context of an individual plate appearance, pitchers and batters respectively attempt to minimize and maximize RE. The straightforward relationship between run creation and wins lends credence to the designation of RE as the outcome variable as opposed to other productive statistics such as OBP or OPS, which have more tenuous connections to wins.

The designation of RE as the outcome variable provides an additional benefit beyond being more representative of the actual objectives of the pitcher-batter interaction; it allows for the valuation and inclusion of non-terminal pitches in the analysis. I assign non-terminal pitches value based solely on whether they are balls or strikes, a strategy grounded on the assumption that the probabilities of success for both pitcher and batter vary across the count. This assumption is regarded as fact in the baseball community; any professional batter will tell you he is more likely to get a base hit on a pitch that is thrown with two balls and no strikes than a pitch that is thrown with no balls and two strikes. Since each count has distinct payoffs associated with it, a pitch can be valued according to its ability to transition from one count to another.

Assigning non-terminal pitches value requires one to calculate the RBOI from each of the twelve possible counts – a straightforward calculation given the proper play-by-play data. Sabermetrician Joe P. Sheehan has calculated these values and makes the count-based RBOI values available on the website 'The Baseball Analysts'. With these values I code each non-terminal pitch with a RE score, which is the difference between the RBOI of the count prior to the pitch and the RBOI of the count resulting from the pitch. For example, since the RBOI from a 1-1 count is -.015 and the RBOI from a 1-2 count is -.082, any pitch thrown for a strike, regardless of whether is a fastball, curveball, slider, or changeup, is coded with a RE of -.067, indicating that, on average, the batter will produce .067 *fewer* runs due to that pitch. Although this method allows for the inclusion of non-terminal pitches in the analysis, it is imperfect, since it neglects any value a pitch may have in setting up another pitch (ie. a high fastball preceding a low changeup), and thus ignores any strategic pitch sequencing on the part of the pitcher. Additionally, all terminal pitches are coded conditional on the count, thus a single (unconditional RE of .49) from a 0-2 count (RBOI of -.104) is coded with a RE of .594. RE scores for all possible terminal and non-terminal pitches are included in the following table:

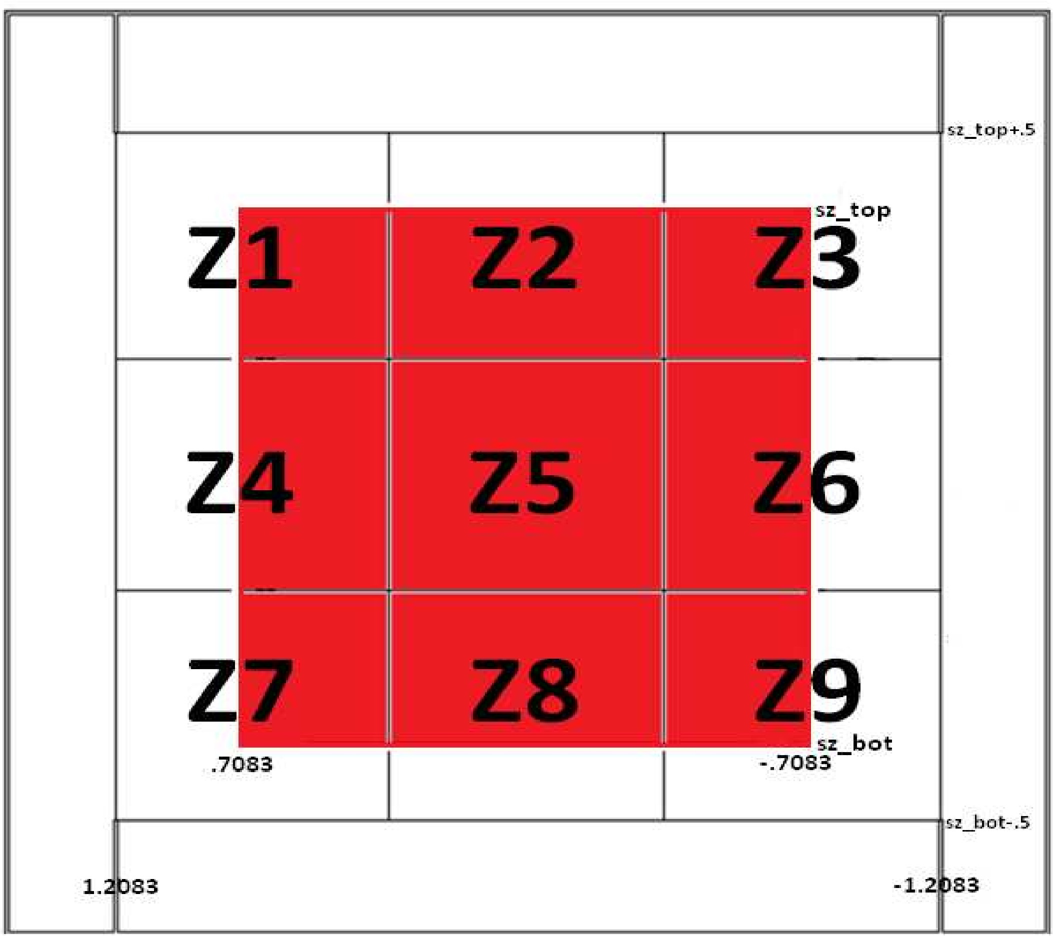
Count	**RE**	**Ball**	**Strike**	**Out**	**Sac**	**Sac Fly**	**HPB**	**1b**	**2b**	**3b**	**HR**
0-0	0.000	0.034	-0.043	-0.289	-0.200	-0.080	0.338	0.494	0.790	1.068	1.407
0-1	-0.043	0.027	-0.062	-0.246	-0.157	-0.037	0.381	0.537	0.832	1.110	1.450
0-2	-0.104	0.022	-0.185	-0.184	-0.096	0.024	0.442	0.598	0.894	1.172	1.511
1-0	0.034	0.063	-0.050	-0.323	-0.234	-0.114	0.304	0.460	0.756	1.034	1.373
1-1	-0.016	0.050	-0.067	-0.273	-0.184	-0.064	0.354	0.510	0.805	1.083	1.423
1-2	-0.083	0.046	-0.206	-0.206	-0.117	0.003	0.421	0.577	0.872	1.150	1.490
2-0	0.097	0.110	-0.062	-0.385	-0.297	-0.177	0.241	0.397	0.693	0.971	1.310
2-1	0.035	0.103	-0.071	-0.323	-0.235	-0.115	0.303	0.459	0.755	1.033	1.372
2-2	-0.037	0.098	-0.252	-0.252	-0.163	-0.043	0.375	0.530	0.826	1.104	1.443
3-0	0.207	0.131	-0.070	-0.496	-0.407	-0.287	0.131	0.287	0.583	0.861	1.200
3-1	0.137	0.201	-0.076	-0.426	-0.337	-0.217	0.201	0.356	0.652	0.930	1.269
3-2	0.062	0.276	-0.351	-0.350	0.262	-0.142	0.276	0.432	0.728	1.006	1.345

The contents of the table indicate that recording a strike is most valuable to a pitcher with two strikes, because this records an out, and the value increases in number of balls due to the fact that more balls are advantageous to the batter because of the increased threat of a BB. Balls are most detrimental to the pitcher when they result in a BB, which can only occur during a count with three balls, and the detriment is increasing in number of strikes since more strikes reduce the likelihood of the batter generating an event that is likely to score a run. Additionally it is clear that terminal pitches are substantially more consequential in terms of RE than nonterminals since they result in events which directly contribute to run production/prevention, while non-terminal pitches can only contribute to run production/prevention through their influence on the count. Although an intentional base-on-balls (IBB) does have a unique RE which is significantly less of that of a BB, it is not included in the preceding table, and all plate-appearances in which an IBB is issued are dropped from the dataset prior to the analysis. The reason for this is that the discretion to issue an IBB is solely awarded to the manager, and thus any plate appearance in which an IBB occurs is not representative of the pitcher's choice of actions.

Controls

Like the Kovash & Levitt study I am able to control for a variety of standard game circumstances such as inning, number of outs, number of runners on base, and count. The detail of the Pitch f/x measurements for each pitch allows me to control for certain factors unobserved in previous studies. Perhaps the most important circumstance that I am able to control for is pitch location. To get a truly ceteris paribus interpretation of the differences in the effectiveness of each pitch type one would ideally compare different pitch types in the same part of the strike zone: clearly a fastball thrown high and inside and a curveball thrown low and away differ in more dimensions than just pitch type. If certain pitches are thrown more frequently in areas of the zone where the batter is at either an advantage or disadvantage, then the estimates of those pitches on RE will be biased.

In order to control for pitch location I divide the entire area over which pitches are located in my dataset, which includes areas both inside and outside of the strike zone, into thirteen different zones. The first nine zones are areas that are likely to have substantially different average REs. For example, if hitters, on average, are able to hit balls better in a certain location, then this location will have a higher average RE. The final four zones encompass areas far outside of the strike zone, and thus pitches in these zones are likely to be mistakes and called balls. How the zones are structured in this area is less important since most of these pitches are called balls and the deviation of average RE across these zones will not vary as drastically as the first nine zones. The following diagram illustrates how I have apportioned the area over which pitches travel into zones for the purpose of location controls.



In the preceding diagram the red area represents the official strike-zone as calibrated by Pitch f/x technicians prior to the start of the game. The area outside of the official strike-zone but within zones 1-9 is a 6-inch wide frame in order to account for umpiring inconsistency. Since umpires are not perfect and often call pitches that narrowly miss the official strike-zone, hitters often swing at these pitches, and thus the potential outcomes and average RE may vary substantially across these areas. Zones 10-13 represent all pitches that are thrown at least 6 inches from any part of the strike zone, which will elicit very few swings and predominantly be called balls. If RE does in fact vary across these zones, and certain pitches are thrown with greater frequency in certain zones, then any estimators that seek to determine the impact of pitch type on RE will be biased if location is not controlled for. Since hitters are generally able to hit pitches in certain locations better than others, including location controls will remove the benefit of a pitch that arises solely as a result of the location and not the pitch type.

Section VI: Results

Replication of the Kovash & Levitt Study

Prior to analyzing the data using the empirical framework set out above I completed a replication of the Kovash & Levitt study, although I run separate regressions for the 2009 and 2010 seasons while they pool the data across all seasons. I replicate the first four estimation procedures set forth in the Kovash & Levitt study; the first three regressions use the exact specifications, while the fourth represents a slight deviation. In their fourth regression they include fixed-effects for pitcher, batter, and pitcher*batter interactions. My strategy is similar, although rather than including fixed-effects for pitcher, batter, and pitcher*batter interactions, which limits the identification to cases in which the same pitcher and batter meet on multiple occasions, I create 'group fixed-effects' and their interactions for pitchers and batters according to characteristics that presumably influence pitcher strategy.

I stratify pitchers across three dimensions: handedness, velocity, and command. The groups for handedness are simple: each pitcher is designated as either right-handed or left-handed. I then create three exhaustive, mutually exclusive groups for both pitcher velocity and command, which allows for classification of each pitcher according to his ability to throw hard and locate pitches. In order to assign pitchers into groups based on velocity I calculate the average fastball start speed for each pitcher on a seasonal basis, and then classify pitchers by the top, middle, and bottom third of these average start speeds. I use a similar strategy for control; for each pitcher I calculate the proportion of balls thrown in 0-0 counts over the span of the season[8], and assign pitchers to groups representing the top, middle, and bottom third for this 'ball rate.'

I also characterize batters across two dimensions: handedness and ability. Like pitchers, batters are classified as either right-handed or left-handed (pitches in which switch hitters are batting are coded according to which side of the batter chooses for that given at-bat). To characterize batters based on ability I calculate the average OPS of each batter on a seasonal basis and assign them to groups representing the top, middle, and bottom third. Although OPS is not a direct measurement of ability, it does capture the batter's performance over a given period of time, and since batter performance is highly visible to all pitchers, it likely exerts strong influence on a pitcher's perception of a given batter's ability.

---

[8] The validity of this measurement of control hinges upon the assumption that pitchers generally attempt to throw a strike on the first pitch of the at-bat.

I include these ‘group fixed-effects’ for both pitchers and batters and their interactions because the groups are based on characteristics of a given pitcher-batter matchup that directly influence a pitcher’s strategy. For example, a right-hander who throws 95 mph with poor control facing an elite left-handed hitter will employ a different strategy than a left-handed slow-throwing command pitcher facing a mediocre right-handed hitter. If it happens that a particular matchup incentivizes the pitcher to throw a certain pitch type more often than others, and that particular matchup is either advantageous or disadvantageous to the hitter (as measured by the differential in average RE from the base-group matchup), then coefficients on the pitch types will be biased.

Including interaction terms for the group fixed-effects (which essentially adds 108 unique ‘matchup’ indicator variables as controls) should emulate a specification which includes pitcher, batter, and pitcher*batter fixed effects as long as the pitcher and batter characteristics that form the basis for the groups represent the essential determinants of pitcher strategy. Doing so will remove the benefit (or penalty) that any pitch receives from being use more (or less) frequently in matchups that favor the pitcher (or batter). In this case the use of broad ‘group fixed-effects’ may be preferable to fixed-effects for pitchers, batters, and their interactions, since it does not limit identification to cases in which the same pitcher and batter meet on multiple occasion, but rather situations in which certain types of pitchers meet certain types of batters on multiple occasions, and thus causes a comparatively smaller increase in standard errors.

The results of the replication are presented in the following table, with each numbered column representing the corresponding specification from the Kovash & Levitt study. Although the general spirit of the results is similar to those found by Kovash & Levitt, some differences are observed, likely due to the separation of datasets by season.

	(1)		(2)		(3)		(4)	
VARIABLES	2009	2010	2009	2010	2009	2010	2009	2010
Fastball	0.093***	0.118***	0.029***	0.047***	0.031***	0.050***	0.049***	0.068***
	-0.009	-0.009	-0.009	-0.008	-0.009	-0.008	-0.009	-0.009
Curveball	-0.159***	-0.082***	-0.036***	0.021*	-0.032***	0.024**	-0.024*	0.032***
	-0.013	-0.012	-0.012	-0.012	-0.012	-0.012	-0.012	-0.012
Slider	-0.088***	-0.051***	-0.032***	0.002	-0.028***	0.006	-0.014	0.020*
	-0.011	-0.011	-0.01	-0.01	-0.01	-0.01	-0.011	-0.011
Inning1					0.066***	0.058***	0.040***	0.029**
					-0.013	-0.013	-0.013	-0.013
Inning2					0.041***	0.038***	0.049***	0.045***
					-0.013	-0.013	-0.013	-0.013
Inning3					0.027**	0.037***	0.036***	0.043***
					-0.013	-0.013	-0.013	-0.013
Inning4					0.062***	0.073***	0.060***	0.069***
					-0.013	-0.013	-0.013	-0.013
Inning5					0.041***	0.054***	0.049***	0.061***
					-0.013	-0.013	-0.013	-0.013
Inning6					0.065***	0.059***	0.061***	0.054***
					-0.013	-0.013	-0.013	-0.013
Inning7					0.022*	0.034***	0.022*	0.034***
					-0.013	-0.013	-0.013	-0.013
Inning8					0.015	0.021	0.012	0.019
					-0.013	-0.013	-0.013	-0.013
Extras					0.008	0.025	0.008	0.026
					-0.023	-0.024	-0.023	-0.024
Nonedown					0.097***	0.095***	0.094***	0.093***
					-0.007	-0.007	-0.007	-0.007
Onedown					0.080***	0.071***	0.080***	0.070***
					-0.007	-0.007	-0.007	-0.007
None_On					-0.052***	-0.053***	-0.054***	-0.058***
					-0.017	-0.017	-0.017	-0.017
One_On					-0.049***	-0.032*	-0.053***	-0.038**
					-0.017	-0.018	-0.017	-0.018
Two_On					-0.036**	-0.021	-0.041**	-0.026
					-0.018	-0.019	-0.018	-0.019
Observations	169,460	158,172	169,460	158,172	169,460	158,172	169,460	158,172
R-squared	0.006	0.005	0.043	0.044	0.045	0.045	0.051	0.051

Column (1) presents the results from the simple regression of OPS on each pitch type with no controls, with changeup serving as the omitted base-group. This specification yields coefficients for fastball, slider, and curveball of .093, -.088, and -.159 for the 2009 season and .118, -.051, and -.082 for the 2010, all of which are individually significant at the 1% level. The magnitude of the coefficient on curveball in the 2009 season is the only substantial deviation from the Levitt & Kovash study for this specification, which indicates that, when nothing else is controlled for, curveballs are more likely to generate a positive outcome for the pitcher in 2009.

The inclusion of count fixed-effects in column (2) greatly reduces the magnitude of the coefficients on all pitch types for both seasons, although the indicators for all pitch types remain significant in the 2009 data while slider is insignificant and curveball is only marginally significant in the 2010 data. The fact that the inclusion of count fixed-effects mutes the OPS gaps across all pitch types suggests that off-speed pitches are thrown with a higher frequency in pitcher's counts and fastballs are thrown with a higher frequency in hitter's counts. This makes sense intuitively as fastballs are easier to locate, and thus pitchers are more likely to rely upon them in situations where the danger of issuing a BB is high (which is always the case in a hitter's count). Kovash & Levitt's results display the same effect upon the inclusion of count fixed-effects.

The inclusion of fixed-effects for inning, outs, and runners on base in column (3) leaves the estimates of interest essentially unchanged, as is the case with the Levitt & Kovash study. The final and most saturated specification in column (4) yields coefficients of .049 and .068 for fastball in the 2009 and 2010 data respectively, which represent increases in magnitude over specification (3). The climbing magnitude on the coefficients for fastball as a result of the inclusion of 'group fixed-effects' and their interactions suggests that on average fastballs are being thrown with a higher frequency in matchups which favor the pitcher, an example of which would be a hard throwing left-hander with excellent control facing a struggling left-handed batsman. The impact of both curveballs and sliders vary greatly across the two seasons: in 2009 slider is insignificant and curveball has a marginally significant positive effect, while in 2010 slider has a marginally significant positive effect and curveball has a highly significant positive effect. Kovash & Levitt report both curveball and slider as insignificant in their fourth specification.

Chow tests for the first three specifications reject the null hypothesis that the effects of all independent variables are identical across the two seasons, though the null hypothesis is not rejected for the fourth specification, indicating that it may be preferable to pool the data for the most saturated specification. It is not surprising that the Chow test for the fourth specification reveals an inability to reject the null hypothesis; the inclusion of over 100 additional regressors (the 'group fixed-effects' and their interactions) substantially reduces the numerator degrees of freedom, thus putting downward pressure on the F-statistic and lowering the probability of a rejection. Although the results of the Chow tests imply that the data should not be pooled for the first three specifications, the following table reports the results of the Kovash & Levitt replication when the seasons are pooled. The table contains results that are remarkably similar to those found in the Kovash & Levitt study, which may suggest that the disparity between the results of the replication and the results of the Kovash & Levitt study is an artifact of running the regressions separately for each season.

VARIABLES	(1)	(2)	(3)	(4)
Fastball	0.107***	0.039***	0.042***	0.060***
	(-0.006)	(-0.006)	(-0.006)	(-0.006)
Curveball	-0.120***	-0.006	-0.003	0.005
	(-0.009)	(-0.009)	(-0.009)	(-0.009)
Slider	-0.068***	-0.014*	-0.009	0.004
	(-0.007)	(-0.007)	(-0.007)	(-0.007)
Inning1			0.062***	0.035***
			(-0.009)	(-0.009)
Inning2			0.040***	0.047***
			(-0.009)	(-0.009)
Inning3			0.032***	0.039***
			(-0.009)	(-0.009)
Inning4			0.067***	0.064***
			(-0.009)	(-0.009)
Inning5			0.048***	0.055***
			(-0.009)	(-0.009)
Inning6			0.062***	0.057***
			(-0.009)	(-0.009)
Inning7			0.028***	0.028***
			(-0.009)	(-0.009)
Inning8			0.018*	0.016*
			(-0.009)	(-0.009)
Extras			0.017	0.017
			(-0.017)	(-0.017)
No_Outs			0.096***	0.094***
			(-0.005)	(-0.005)
One_Out			0.076***	0.075***
			(-0.005)	(-0.005)
None_On			-0.052***	-0.056***
			(-0.012)	(-0.012)
One_On			-0.041***	-0.046***
			(-0.012)	(-0.012)
Two_On			-0.029**	-0.034***
			(-0.013)	(-0.013)
Observations	327,632	327,632	327,632	327,632
R-squared	0.005	0.043	0.045	0.05

Results from an Alternative Empirical Strategy

The following section presents the results from the empirical strategy set forth in Section VI of this paper, in which RE is the dependent variable. This strategy analyzes all pitches from each plate appearance, rather than just the first, as in the Weinstein-Gould study, or the last, as in the Kovash & Levitt study. Although the effectiveness of pitch types differ across seasons, as is the case in the Kovash & Levitt replication above, the general tone the results are strikingly similar to those produced by both Weinstein-Gould and Kovash & Levitt; not only do pitchers not optimize across pitch types, but they appear to rely too heavily upon the fastball over both seasons analyzed. The results for both seasons across all functional forms are displayed in the table on the following page.

Column (1) presents the estimates from a simple regression of RE on fastball, curveball, and slider, with changeup serving as the omitted base-group. The disparity across seasons is again readily apparent. The 2009 data generates a coefficient of 0.000 for fastball, while both curveball and slider have similarly negative impacts on RE and are both significant at the 1% level. The 2010 data paints an extremely different picture: both fastball and slider are statistically significant at the 1% level with coefficients of .003 & -.004 respectively, while curveball has a coefficient of 0.000. Since this specification includes no controls whatsoever, the results merely indicate the difference between the average RE of each pitch type and the base-group, and thus are not especially useful due to the fact that they neglect to account for the impact on RE of a variety of game situations.

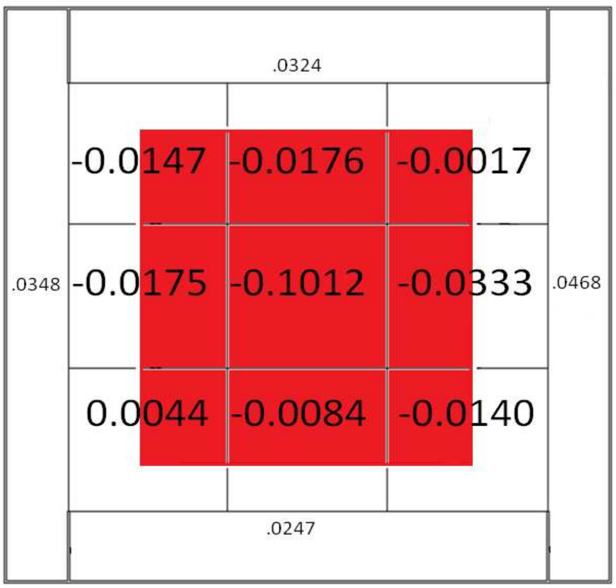
	(1)		(2)		(3)		(4)	
VARIABLES	2009	2010	2009	2010	2009	2010	2009	2010
Fastball	0.000	0.003***	0.004***	0.007***	0.004***	0.007***	0.006***	0.009***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Slider	-0.008***	-0.004***	-0.008***	-0.003***	-0.007***	-0.003**	-0.006***	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Curveball	-0.007***	0.000	-0.008***	-0.001	-0.008***	0.000	-0.007***	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Inning1					0.009***	0.008***	0.005***	0.004***
					(0.001)	(0.001)	(0.001)	(0.001)
Inning2					0.006***	0.005***	0.006***	0.005***
					(0.001)	(0.001)	(0.001)	(0.001)
Inning3					0.006***	0.006***	0.006***	0.006***
					(0.001)	(0.001)	(0.001)	(0.001)
Inning4					0.009***	0.009***	0.008***	0.008***
					(0.001)	(0.001)	(0.001)	(0.001)
Inning5					0.007***	0.008***	0.007***	0.008***
					(0.001)	(0.001)	(0.001)	(0.001)
Inning6					0.010***	0.009***	0.008***	0.007***
					(0.001)	(0.001)	(0.001)	(0.001)
Inning7					0.005***	0.005***	0.004***	0.005***
					(0.001)	(0.001)	(0.001)	(0.001)
Inning8					0.003**	0.003**	0.003*	0.003**
					(0.001)	(0.001)	(0.001)	(0.001)
Extras					0.003	0.004*	0.002	0.004*
					(0.002)	(0.002)	(0.002)	(0.002)
One_Out					-0.002***	-0.003***	-0.002**	-0.003***
					(0.001)	(0.001)	(0.001)	(0.001)
Two_Outs					-0.012***	-0.011***	-0.011***	-0.011***
					(0.001)	(0.001)	(0.001)	(0.001)
One_On					0.003***	0.005***	0.002***	0.004***
					(0.001)	(0.001)	(0.001)	(0.001)
Two_On					0.004***	0.006***	0.004***	0.006***
					(0.001)	(0.001)	(0.001)	(0.001)
Three_On					0.008***	0.009***	0.008***	0.009***
					(0.002)	(0.002)	(0.002)	(0.002)
Observations	599,793	562,011	599,793	562,011	599,793	562,011	599,793	562,011
R-squared	0	0	0.013	0.013	0.013	0.013	0.016	0.016

Specification (2), which adds fixed-effects for pitch location, illustrates a substantial magnifying effect for the fastball coefficient across both seasons. The coefficients on the fastball indicator increase from 0.000 to 0.004 in the 2009 data and 0.003 to .007 in the 2010 data, implying that the location-specific benefits to the fastball were relatively constant across the seasons. The impact of throwing a fastball on RE becomes statistically significant at the 1% level across both seasons, whereas in specification (1) fastball had no effect (significant or practical) on RE in the absence of location controls in the 2009 data. The inclusion of location controls has virtually no effect on the impact of throwing curveballs and sliders, with the coefficients on these indicators remaining constant across the first two specifications in both seasons.

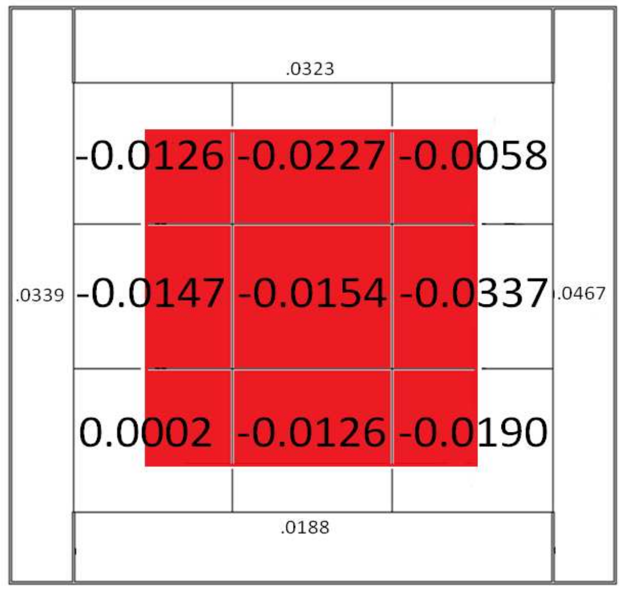
The upward sensitivity of the impact of throwing a fastball on RE suggests that, on average, fastballs are thrown in locations in which hitters are less likely to produce damaging outcomes for the pitcher. This result is expected as fastballs are considerably easier for pitchers to command than off-speed pitches, and thus are more readily able to exploit location-specific batter disadvantages. Since, in the context of this study, the pitcher-batter interaction is modeled such that the pitcher only chooses pitch types and not location, including fixed-effects for pitch location strips away the benefit a pitcher receives from strategically locating fastballs. It is also unsurprising that the inclusion of controls for pitch location does not affect the impact of throwing off-speed pitches since curveballs and sliders are notoriously difficult to locate effectively, and as a result are unlikely candidates for pitches to be strategically located.

The following tables illustrate the average RE for each location zone across the two seasons, as well as the proportion of fastballs to total pitches thrown in each location zone. As expected from the upward sensitivity of the fastball coefficient to the inclusion of the location controls, the tables indicate that pitchers tend to throw fewer fastballs in zones in which the hitter is at a substantial advantage.

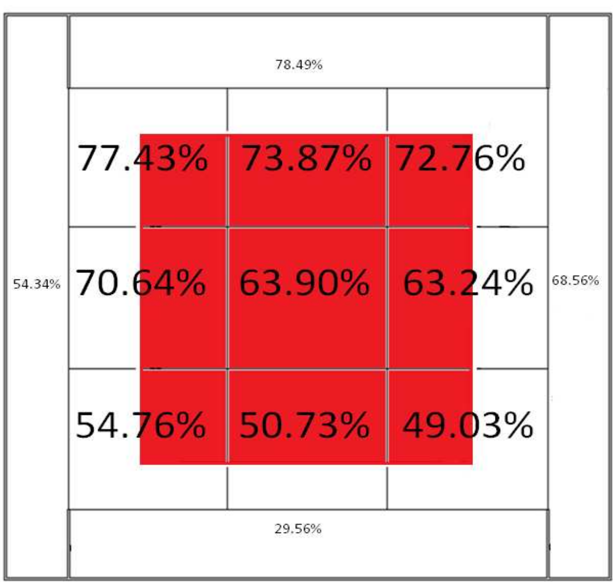
Average RE by Zone (2009)



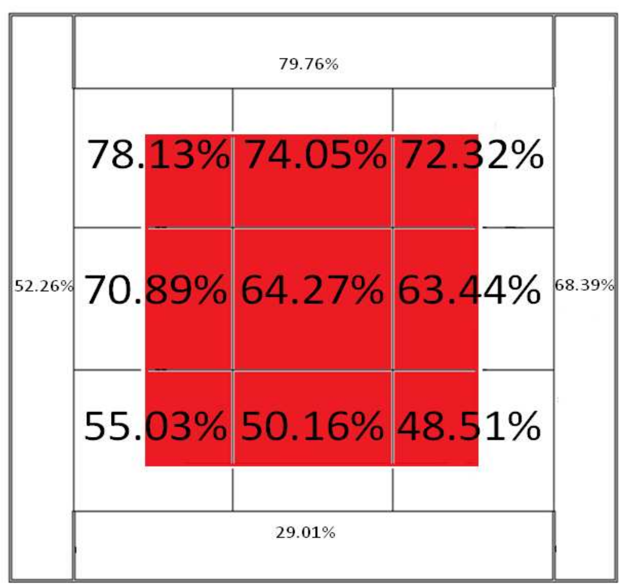
Average RE by Zone (2010)



Fastball Rate by Zone (2009)



Fastball Rate by Zone (2010)



The preceding tables indicate that, amongst pitches located within Zones 1-9 (those zones which are likely to elicit swings), batter leverage is greatest on low pitches. Additionally, the proportion of fastballs to total pitches thrown in Zones 1-9 is lowest on low pitches, suggesting that pitchers are able to avoid locations in which they are most vulnerable to offensive damage with fastballs. The consistently high average RE values for Zones 10-13 is due to the fact that not only are these pitches more likely to be called balls (which invariably carry positive RE values), but also that they are much less likely to elicit swings than pitches in Zones 1-9, and thus rarely generate outs, which carries the highest negative RE value. Pitchers also seem to avoid these zones with fastballs, especially Zones 10 and 13 which represent the outside and low locations, although admittedly to a lesser extent than they do with low and hittable locations. The above tables clearly demonstrate that the fastball provides a benefit to pitchers in the fact that it is easier to throw with precision, which allows pitchers to strategically locate it. This effect gives the indicator for fastball a downward bias in the specification (1), which is corrected for with the inclusion of location controls.

In accordance with the Kovash & Levitt replication, the inclusion of fixed-effects for inning, outs, and number of runners on base, which corresponds to column (3), produces virtually no effect on the variables of interest. This suggests that no pitch type is thrown substantially more or less often in a game situation in which the batter is a clear advantage and disadvantage. Column (4) adds the ‘group fixed-effects’ explained above (although in this case batters are placed into groups segmented by average RE, not OPS), and has a generally magnifying effect on the impact of a fastball. The coefficient on fastball rises from .005 to .007 in 2009 and .008 to .010 in 2010, with both estimates remaining statistically significant at the 1%

level. In 2009 both curveball and slider are relatively unaffected by the inclusion of 'group fixed-effects'; each has a coefficient of -.006 and is significant at the 1% level. In the 2010 data the inclusion of 'group fixed-effects' moderately reduces the impact of throwing a slider, with coefficient falling from -.003 in specification (3) to -.001 in specification (4), rendering is statistically insignificant.

The results from the most saturated specification indicate that, not only are pitchers not optimizing, but they employ different sub-optimal strategies across seasons. The 2009 results suggest that pitchers threw too many fastballs and too few curveballs and sliders, relative to changeups. The 2010 results seem to indicate that pitchers adopted a relatively better strategy than in 2009 by throwing curveballs and sliders such that the marginal benefit from throwing them was reduced to zero, but continued to rely too heavily on the fastball, which resulted in a higher penalty to throwing the fastball. The differences in the estimates of all regressors across both seasons are statistically significant at the 1% level across all specifications, with Chow-tests for each specification yielding F-statistics in excess of the associated critical values.

The estimates from specification (4) which uses RE as the dependent variable indicate that a pitching staff could have allowed 17 and 24 fewer runs in the 2009 and 2010 seasons respectively, by throwing 10% fewer fastballs, though this calculation assumes that each fastball is replaced with a changeup, and that the RE differential across pitch type remains constant despite the change in strategy. These estimates likely overstate the downside to throwing too many fastballs, as replacing fastballs with changeups (or any other pitch for that matter) would likely decrease the RE gap between the fastball and the replacement pitch due to behavioral responses by the hitter. Despite the substantial adjustments to the empirical strategy, my results are surprisingly similar to those of Kovash & Levitt, who estimate that a pitching staff could reduce the numbers of runs allowed over a season by approximately 15, though the authors do not specify which pitches (and in what proportions) should replace the fastballs. Clearly the ‘runs added’ estimates are highly dependent upon the specifications from which we draw the estimates, so I include the following table, which contains ‘potential runs saved’ estimates for each specification across both seasons, in order to quantify the sensitivity of these estimates to inclusion of various controls.

‘Potential Runs Saved’ by Season and Specification

Season	(1)	(2)	(3)	(4)
2009	0.00	9.72	9.72	14.58
2010	7.29	17.01	17.01	21.87

Sensitivity of the Estimates to Subsets of the Data

The estimates generated above suggest that MLB pitchers fail to optimize across pitch types with respect to RE, and that they do so to an extent that inflicts tremendous losses on their teams, and perhaps themselves individually in the form of less lucrative contracts than they could otherwise command. This result is highly counter-intuitive given that professional pitchers are expert-level agents; possess extensive information concerning the strengths, weaknesses, and behavioral patterns of their opponents; and face tremendous monetary consequences for sub-optimal performance. If the empirical strategy presented above correctly models the pitcher-batter interaction so that the that failure to detect optimal mixed-strategies is not merely an artifact of improper characterization of the interaction, then a logical next step is to determine what factors cause pitchers to deviate from optimal play.

Isolating instances in which pitchers adhere to the predictions of the theory will allow for the identification of characteristics that systemically differ across those who optimize and those who do not, and may contain some predictive power regarding optimization. The table on the following page presents the estimates from specification (4) for different subsets of the 2009 data in order to illustrate a first-attempt at identifying subsets of the data in which optimization is observed. The first three columns represent the bottom, middle, and top third of pitchers by average RE, while the last the three columns represent the bottom, middle, and top third of batters by average RE.

The table reveals an unexpected trend: the positive RE gap on fastball is actually increasing in pitcher performance, with the best pitchers receiving the worst outcomes on the fastball. On the other hand the negative impact of off-speed pitches on RE diminishes with pitcher performance, implying that the best pitchers tend to throw curveballs and sliders such that the resulting RE is equalize with that of the changeup. While these results are initially surprising, one potential explanation lies in the failure of my empirical strategy to account for pitch-sequencing concerns. As stated in the Section V, I value non-terminal pitches solely based upon their classification as balls or strikes. In this case, a pitch that is used strategically to set up the batter for failure on the following pitch obtains no value for its role in recording the out, but rather only for its role in transitioning the count. It may be that the best pitchers in MLB are able to effectively retire batters due to sophisticated pitch sequencing. If these pitchers primarily use fastballs to set up batters, and then record outs with off-speed pitches, the fastball will generate fewer negative RE values, and thus will look as though it results in a worse outcome on average.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Fastball	0.004*	0.007***	0.010***	0.009**	0.006***	0.006***
	(0.002)	(0.001)	(0.002)	(0.004)	(0.001)	(0.001)
Slider	-0.010***	-0.004**	-0.003*	-0.013**	-0.006***	-0.006***
	(0.003)	(0.002)	(0.002)	(0.005)	(0.002)	(0.002)
Curveball	-0.007**	-0.007***	-0.001	-0.022***	-0.007***	-0.006***
	(0.003)	(0.002)	(0.002)	(0.006)	(0.002)	(0.002)
Inning1	-0.003	0.002	0.003	0.010	0.001	0.006***
	(0.004)	(0.002)	(0.002)	(0.009)	(0.002)	(0.002)
Inning2	0.001	0.003	0.005**	-0.005	0.006***	0.007***
	(0.004)	(0.002)	(0.002)	(0.006)	(0.002)	(0.002)
Inning3	0.001	0.003	0.004**	-0.012*	0.005**	0.007***
	(0.004)	(0.002)	(0.002)	(0.006)	(0.002)	(0.002)
Inning4	0.006*	0.005**	0.005**	-0.012*	0.009***	0.009***
	(0.004)	(0.002)	(0.002)	(0.006)	(0.002)	(0.002)
Inning5	0.004	0.004	0.006***	-0.005	0.005**	0.008***
	(0.004)	(0.002)	(0.002)	(0.006)	(0.002)	(0.002)
Inning6	0.009**	0.005**	0.006***	-0.005	0.010***	0.008***
	(0.004)	(0.002)	(0.002)	(0.007)	(0.002)	(0.002)
Inning7	0.005	0.002	0.004**	-0.010	0.003	0.006***
	(0.004)	(0.002)	(0.002)	(0.007)	(0.002)	(0.002)
Inning8	0.005	0.000	0.002	-0.002	0.003	0.003
	(0.004)	(0.002)	(0.002)	(0.007)	(0.002)	(0.002)
Extras	0.009	0.001	-0.000	-0.007	-0.001	0.005
	(0.007)	(0.004)	(0.003)	(0.012)	(0.004)	(0.003)
Onedown	-0.001	-0.001	-0.003**	-0.003	-0.002	-0.002**
	(0.002)	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)
Twodown	-0.014***	-0.010***	-0.009***	-0.004	-0.010***	-0.012***
	(0.002)	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)
One_on	0.006***	0.000	0.002*	-0.007**	0.003**	0.003***
	(0.002)	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)
Two_on	0.011***	0.002	0.001	0.001	0.004***	0.004***
	(0.002)	(0.001)	(0.001)	(0.004)	(0.001)	(0.001)
Three_on	0.013***	0.004	0.008***	-0.002	0.009***	0.008***
	(0.004)	(0.003)	(0.003)	(0.006)	(0.003)	(0.003)
Observations	113,746	263,241	222,806	19,975	225,102	354,716
R-squared	0.012	0.015	0.021	0.057	0.017	0.011

Section VII: Conclusions

The results of this study are in direct agreement with both papers that have tested for optimal mixed-strategies in the pitcher-batter interaction: MLB pitchers appear to rely too heavily upon the fastball, rendering it significantly less effective than the changeup, curveball, and slider. The increased impact of throwing a fastball on RE in the presence of controls for pitch location suggests that the estimates on the impact of a fastball in the previous studies, which were unable to control for pitch location, may have contained a downward bias, perhaps causing the authors to understate the penalty inflicted upon pitchers from throwing fastballs too frequently. The impact of throwing both sliders and curveballs on RE varies substantially across seasons and specifications, thus I can draw no strong conclusions regarding their effectiveness compared to the changeup.

The consistently positive impact of throwing a fastball on RE across all specifications used in all studies analyzing this topic serves as strong evidence that MLB pitchers do not optimize according to the predictions of the minimax theorem, and thus perform below their potential. Since compensation in professional sports is a direct function of performance, it seems that players are 'leaving money on the table'. Interesting work on this issue going forward will attempt to identify groups of MLB pitchers whose strategies do adhere to the standard predictions of the minimax theorem. Doing so could shed light upon what traits, if any, systematically differentiate those that optimize from those that do not. Locating specific characteristics or situations in which pitchers do optimize will provide more general behavioral insight into why certain agents underperform their potential in settings outside of professional sports which been unable to generate data indicating the use of optimal mixed-strategies.

Another potential explanation for the inability to detect optimal mixed-strategies is that the studies analyzing this issue have not successfully accounted for the nuances of the pitcher-batter interaction in their empirical strategies. The finding in Section VI that the best pitchers appear to employ the worst strategies (with respect to overuse of the fastball) may be interpreted as evidence that future work on this topic must attempt to incorporate a valuation system for strategic pitch sequencing. In this case the benefit (or penalty) to a terminal pitch that records an out (or results in an unfavorable outcome to the pitcher) must be somehow apportioned over previous pitches which may have contributed to this final outcome. The use of RE as the outcome variable in this study represents a first step towards this end in that is able to assign value to non-terminal pitches, though it is clearly limited in its ability to account for strategic pitch sequencing. Future studies attempting to identify optimal mixed-strategies amongst MLB pitchers can produce more credible estimates by incorporating a means of attributing value to pitches for their use in strategic pitch-sequencing situations.

Section VIII: References

Birnbaum, Phil. "Don't Use Regression to Calculate Linear Weights." *Sabermetric Research*. 27 Oct. 2009. Web. Dec. 2010. <http://sabermetricresearch.blogspot.com/2009/10/dont-use-regression-to-calculate-linear.html>.

Byous, Jon. "Play Ball!" *Oracle: Sun Developer Network*. Nov. 2003. Web. Nov. 2010. <http://java.sun.com/developer/technicalArticles/ThirdParty/Tendu/>.

Chiappori, P.A., S. Levitt, and T. Groseclose. 2002. Testing mixed strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer. *American Economic Review* 92 (4): 1138 – 1151.

DiFucci, Dom. "What Is Sabermetrics?" *Mets Chronicle*. 31 Oct. 2010. Web. Apr. 2011.

Klaassen, Matt. "Custom WOBA and Linear Weights Through 2010." *Beyond the Box Score*. 4 Jan. 2011. Web. Mar. 2011. <http://www.beyondtheboxscore.com/2011/1/4/1912914/custom-woba-and-linear-weights-through-2010-baseball-databank-data>.<http://www.metschronicle.com/2010_10_01_archive.html>.

Kovash, Kenneth & Levitt, Steven D. (2009) Professionals Do Not Play Minimax: Evidence from Major League Baseball and the National Football League. NBER Working Paper No. w15347.

McCauley, Janie, Steven Wine, Howard Fendrich, Rick Gano, Mark Didtler, Ian Harrison, Pete Kerzel, and Steve Herrick. "Major League Baseball Teams Now Scouting Umpires Too." *SI LIVE*. 11 May 2010. Web. Feb. 2011. <http://www.silive.com/sports/index.ssf/2010/05/major_league_baseball_teams_no.html>.

Nathan, Alan. "The Sportsvision Pitch Tracking System (Pitch F/x) How Does It Work?" *The Physics of Baseball*. 3 Aug. 2007. Web. 20 Aug. 2010. <http://webusers.npl.illinois.edu/~a-nathan/pob/technology.htm>.

Nathan, Alan. "Tracking Pitches Using Video Technology." *The Physics of Baseball*. 27 Mar. 2010. Web. 20 Aug. 2010. <http://webusers.npl.illinois.edu/~a-nathan/pob/pitchtracker.html>.

Newman, Mark. "Pitch F/x Adds Insight to Watching Games." *The Official Site of Major League Baseball*. 2 Oct. 2007. Web. 16 Apr. 2011. <http://mlb.mlb.com/news/article.jsp?ymd=20071002>.

"Official Rules." *The Official Site of Major League Baseball*. Web. 15 Jan. 2011. <http://mlb.mlb.com/mlb/official_info/official_rules/foreword.jsp>.

Palacios-Huerta, Ignacio. (2003). Professionals Play Minimax. *Review of Economic Studies* 70: 395-415.

Sheehan, Joe P. "More Run Values." *The Baseball Analysts*. 21 Feb. 2008. Web. Jan. 2011. <http://baseballanalysts.com/archives/2008/02/writing_about_t.php>.

"Sportvision: Baseball." *Sportvision: Changing the Game*. Web. 20 Aug. 2010. <http://www.sportvision.com/base-pitchfx.html>.

Tango, Tom. "Game Theory on Pitch Selection." *THE BOOK*. 28 Sept. 2009. Web. Oct. 2010. <http://www.insidethebook.com/ee/index.php/site/comments/game_theory_on_pitch_selection/>.

Tango, Tom. "Run Values of Events." *THE BOOK*. 9 Feb. 2010. Web. 16 Feb. 2011. <http://www.insidethebook.com/ee/index.php/site/comments/run_values_of_events/>.

Walker, M., and J. Wooders. (2001). Minimax play at Wimbledon. *American Economic Review* 91 (5): 1521-38.

Weinstein-Gould, Jesse. (2009). Keeping the Hitter Off Balance: Mixed Strategies in Baseball. *Journal of Quantitative Analysis in Sports* 5 (2).