Discussion Papers In Economics And Business

Pitch Call Discrimination in Major League Baseball:

The Effect on the Observed Performance and the Salaries

Reio Tanji

Discussion Paper 22-02

August 2022

Graduate School of Economics
Osaka University, Toyonaka, Osaka 560-0043, JAPAN

Pitch Call Discrimination in Major League Baseball: The Effect on the Observed Performance and the Salaries

Reio Tanji[†]

Abstract

This paper identifies discrimination in the professional baseball league in the United States. We consider players, umpires, and team managers as employees, middle managers, and employers in general workplaces. Using huge pitch-by-pitch tracking data of the Major League Baseball reveals that umpires from North America favor players with the same region when a pitcher from North America and a batter from other regions are facing. The impact of this is dramatic: players from other regions lose their chances to hit by the unfair pitch call, which values to loss of about $130,000 for in three years.

■JEL Classification: D91, J01

■keywords: sports, discrimination, in-group bias, baseball

[†] Osaka University, Graduate School of Economics. email:u280537h@ecs.osaka.u-ac.jp

1 Introduction

Labor economics studies have long examined discrimination and have classified the favoritism exhibited by evaluators toward individuals belonging to the same group as themselves, as in-group bias. In-group bias has severe implications for the labor market.

This study presents an empirical analysis of in-group bias using a large dataset from Major League Baseball (MLB). Situations there are common in the US even in the present day, with regular reports of incidents of discrimination. Thus, understanding the extent of discrimination in the MLB is essential for visualizing the existence of discrimination in society as a whole.

Specifically, we use a radar-based high-precision ballistic measurement device called the *Trackman*, which was installed in all franchise ballparks of MLB teams from 2015 to 2019 to record and disclose detailed information on the pitchers' thrown ball trajectory and the batters' batted ball launch angle and launch velocity. This study focuses on the pitch location as it passes the home plate. When the pitcher throws a pitch and the batter does not hit it (takes), the plate umpire judges whether it is a strike or ball. Although there is a strict regulation where a pitch should be called a strike, the umpire's decision is visually validated. Thus, two pitches passing through an exact location can be given different calls, especially near the edge of the strike zone. This study considers this system can identify the objective information of the pitches. By comparing this to the actual pitch calls, an evaluation by the plate umpire, we can examine how their demographic characteristics affect the evaluator's (plate umpire) assessment of the achievement of the workers (players).

Because we can use workers' (athletes') performance as an objective measure, and because it constitutes a large dataset, there has been a large body of literature on discrimination using sports data. Among these, the most relevant to this paper is Parsons et al. (2011). They use the MLB tracking data to analyze the impact of racial discrimination by the plate umpire, as in this paper. The most important advantage of this dataset is the random assignment of umpires and almost external variation in installations of "*QuesTec*."[1] They found that racial differences in pitchers and umpires increased the rate of strike calls. Tainsky et al. (2015) also conducted a follow-up study using the same data.

Compared to the previous studies, we emphasize the following new contributions: First, we focus on how umpires' bias distorts the players' performance statistics; we also quantify the magnitude of the effect on the salary players receive. Objective information about pitch locations enables us to see the probability that a specified pitch should be called strike and to

[1] "*QuesTec*" is a technology installed in some of the franchise stadiums of the MLB teams, to provide feedback to umpires on the accuracy of their decisions.

compare this result with the actual call. This, in turn, allows us to argue about the size of the impact umpires have through pitch calls. Notably, this bias may distort the performance data of discriminated players, making their performance appear worse than it is in reality. This negative effect is particularly important, because general managers offer contracts based on observed performance indicators. Hence, players discriminated against by umpires may be forced to sign contracts on unfavorable terms/contracts of a lower monetary value than those favored.

To explore this problem this study takes advantage of a video-based judgment review system called "Manager Challenge." Using this system, we can capture almost all indirect negative impacts of umpire bias on player compensation. This study is unique as it deals with two sources of discrimination in the same context. One is the discrimination exhibited by umpires(evaluators in the workplace), and the other is by managers (principals) who offer contracts to players (player contracts). In a typical workplace, middle managers evaluate worker performance. Hence, based on their reports, company heads decide whether to renew the contract with workers. This environment is similar to the example we address in this study and provides suggestive results for distinguishing between discrimination by middle managers and discrimination by the president. From this perspective, this study also has implications for the more general context of labor economics.

Second, this study provides a more evidence-based discussion of the endogeneity problem implied by pitch tracking data, which is suggested in the prior literature. For example, pitchers discriminated against may avoid pitching to the edge of the strike zone. Moreover, batters may swing more aggressively, and we cannot observe such pitches as ones to be called. If so, biased calls may result in worse performance for batted balls hit by minority players. This study explores this concern by analyzing the behavioral responses of batters toward pitches.

Our main results are as follows. First, the pitch calls made by umpires from the United States (US) suggest the influence of a nationality bias, based on batter nationality. Specifically, in situations where the pitcher is from the same country as the umpire but the batter is not, the umpire calls strikes at a higher rate (i.e., decisions are largely made in favor of the pitcher). Based on the average probability of a strike call estimated from the information about the two-dimensional location of the pitch, the variance of the probability in such a situation is about 0.3 percentage points per taken pitch.

The second result attempts to determine the monetary cost of nationality bias. Based on the first result, minority batters who consistently have their names in the lineup are called about 38 more strikes per year. This worsens the performance statistics that the managers observe, resulting in a loss of approximately $40,000 in compensation per year that they could have received.

The third result concerns endogeneity in players' decision-making. Discrimination by plate umpires may cause behavioral changes in players who perceive it, but the analysis conducted in this study did not provide strong evidence to support these behavioral changes.

The remainder of this paper is organized as follows. Section 2 reviews the literature on discrimination, especially empirical literature in the sports context. Section 3 explains the institutional background of baseball or MLB and the decision making of plate umpires in pitch calls. Section 4 discusses the empirical data. Section 5 presents the empirical model used for identification, and Section 6 provides the results. Section 7 discusses the results, and Section 8 concludes the study.

2 Literature Review

Owing to the availability of rich datasets, many empirical studies use the MLB context for research on in-group or racial bias. Scully (1974), who conducted one of the oldest and most well-known studies on the topic, estimated the profit function of MLB teams and analyzed the characteristics of the players and the team's revenue, proxied by the number of attendees. They highlighted that an increase in the percentage of African American players in the team's roster harms the team's revenue, and also discussed consumer bias by the audience. A similar study by Hill et al. (1982) demonstrated that, if the home team's starting pitcher announced in advance is a minority pitcher, attendance for that game decreases. Nardinelli and Simon (1990) examined the MLB player card market and found that non-white pitchers with comparable performance were valued about 13% lower than white pitchers (10% lower for non-pitchers).

Another target of research on discrimination using data from professional sports is discrimination by teams offering contracts to players. We can observe the quantified contribution of each player to the team through performance statistics, such as the number of homeruns or strikeouts. Moreover, we can see the amount of compensation each player receives (i.e., the evaluation by the general managers). This situation allows us to argue whether the salary for a player is adequate or not. Many studies test the existence of racial discrimination by team managers. Krautmann et al. (2000) regarded performance statistics as performance indicators and showed that non-white players are more cost-effective compared to white players. Essentially, teams can acquire non-white players through cheaper contracts. Non-white players are discriminated against by companies (teams) in the player labor market. Similarly, Holmes (2011) conducted analyses on discrimination by estimating the relationship between player performance indicators and compensation.

3 Background

3.1 Pitch Call in the MLB

First, we describe the process of judging balls and strikes. Hereafter, we refer to this process as "pitch calls."

In MLB, typically, there are four umpires: the plate umpire and the respective base umpires at the first, second, and third bases. The four umpires comprise a unit called a "crew," As a rule, the league assigns one crew to each game in the regular season. An umpire does not specialize in any single position but rotates across all four positions sequentially. Hence, an umpire will be a plate umpire in approximately one out of every four games. Many umpires are graduates of technical schools and, as players, work their way up through ranks in lower categories (Minor League Baseball, or MiLB), to earn MLB contracts. Salaries paid to umpires are comparable to those paid to highly skilled professionals[2]. Because their decisions are televised throughout the US and worldwide, they have incentives to make accurate and unbiased decisions.

Making pitch calls is one of the plate umpire's most important tasks. If the batter takes (does not swing) the pitch, the plate umpire calls "ball" or "strike" based on the pitch location as it crosses home plate. The strike zone is defined as follows:

> The official strike zone is the area over home plate from the midpoint between a batter's shoulders and the top of the uniform pants – when the batter is in his stance and prepared to swing at a pitched ball – and a point just below the kneecap. In order to get a strike call, part of the ball must cross over part of home plate while in the aforementioned area.

quoted from MLB.com: Strike Zone

In each plate appearance, if an umpire calls the third strike (caught by the catcher) before the batter hits the pitch into fair territory, the batter is out (strikeout). However, if four pitches pass outside the strike zone and the batter takes them, he is entitled to move to first base. The number of balls and strikes during the at-bat (pitch count) also significantly affects the result of each plate appearance (see Table 1). Moreover, whether an individual pitch is called a strike or a ball is a pivotal issue in player performance (see Table 1).

Although the strike zone is well-defined, not all decisions follow it. Figure 1 illustrates the physical pitch location and actual decision. The bold rectangle in the figure shows the strike zone according to the definition—any pitch that passes inside this zone should be called a strike.

[2] For more detailed information about MLB umpires, see Much required to become MLB umpire

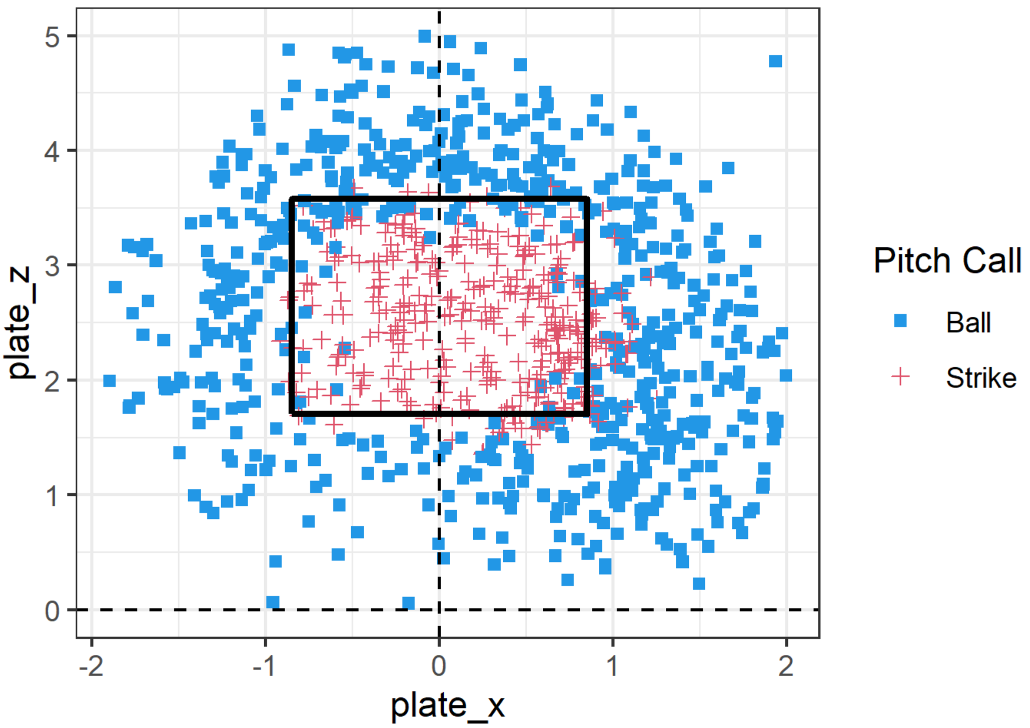
Table 1 wOBA by Pitch Counts

Balls \ Strikes	0	1	2
0	.406	.372	.173
1	.413	.390	.186
2	.430	.407	.210
3	.678	.580	.389

wOBA (the abbreviation for Weighted On-Base Average) is one of the most important and popular statistics to evaluate batters. See https://library.fangraphs.com/offense/woba/ for the details.

However, in practice, a pitch that passes through the strike zone can be called a ball and vice versa. The plate umpire judges each pitch visually; this judgment cannot be overturned once it has been called. Therefore, if the umpire makes a discriminatory decision (regardless of whether the tendency occurs subconsciously), the effect will be to the detriment of minorities in the form of a lower performance index because of an unfavorable decision.

Figure 1 Pitch location and Pitch calls



Batter: Mike Trout, 1,000 pitches randomly picked up.

3.2 Pitch Tracking System in the MLB

We will now explain how MLB acquires pitch-tracking data in each game and how these data are utilized in empirical analyses. Tracking data include physical information on players and balls using cameras and radar technology. This system accumulates pitch-by-pitch data and the results of each play, with tags for players and games. In MLB, the velocity and rotation of pitches and the two-dimensional position of the ball as it passes the home plate are recorded for each pitch (i.e., pitch location), along with other information on the game. In each game, information on about 300 pitches is accumulated. The MLB regular season consists of approximately 2,400 games [3], so we obtained data on approximately 720,000 pitches for every season.

In 2004, the first MLB team started to install equipment to acquire tracking data. This system, called *Pitch F/X*, uses cameras to capture the trajectory of each pitch and movement produced by ball rotation. Because this dataset provides objective information on pitch locations, they began to use it to provide feedback on the pitch calls. In Parsons et al. (2011), the authors exploited the exogenous differences in the presence of this feedback system, called *QuesTec*.[4]. In 2008, all 30 MLB teams installed the *Pitch F/X* system in their franchise ballparks. Since then, the dataset has been available in Baseball Savant.

In 2015, another tracking system, *TrackMan*[5], replaced *Pitch F/X*. This new instrument uses Doppler radar to capture balls, instead of using high-speed cameras, making the capture of the exit velocity or exit angle of the batted balls possible, in addition to pitches. This system allows us to more objectively estimate the quality of batted balls. Based on the physical information about the ball, we can quantify the probability of whether the ball will be a hit or home run. Using sabermetrics, we can also calculate the number of runs the batted ball creates. All the MLB ballparks were installed with *TrackMan* at the beginning of the 2015 season; hence, data are available for almost all the games played each year[6]. Tagged by information of the games and players, data tracked by the radar constitute a vast dataset. This dataset includes more detailed information, such as defensive shifts recorded by high-precision cameras or the efficiency of

[3] The regular-season game is held to determine the champions of the six divisions (two leagues $\times$ three divisions). Each team plays 162 games and, based on the results, their right to enter postseason games to compete with the all-league champion, called the World champion, is determined.

[4] In ballparks equipped with *QuesTec*, every pitch call is recorded and evaluated by the League. Conversely, those without the system do not observe the actual locations of the pitches; hence, umpires may feel less monitored, which may lead to more biased judgments.

[5] The *Pitch F/X* system continued until the 2020 season; however, the official version of the dataset was tracked via *TrackMan*.

[6] Most of the official MLB games are held at MLB franchise ballparks. Only a few games have taken place outside the US (such as in Japan, Mexico, or China). Portable devices enable the acquisition of data even in stadiums without *TrackMan*, as in the United Kingdom in 2019.

routes used by fielders to track batted balls[7]. This system is called Statcast.

3.3 Video-Based Judge Review System

Another essential institutional background for the analysis in this study is the introduction of a replay system. At the start of each game in the regular season, the managers of both teams receive a one-time right to request a replay review of close plays. If they exercise the right in the short interval before the next play, a video replay of the previous play is reviewed. The umpires cannot refuse to entertain a review request. If the coach's request is right, the judge is overturned. Currently, each team keeps staff outside the dugout to review the video, and they are responsible for determining whether the play should be challenged and for informing the coach. Additionally, the umpires may also request a review when they regard it as necessary.

In all cases, it is not the umpires in charge of each game that reviews the judge but umpires operating in a specialized facility called the "Replay Command Center" in New York. They scrutinize the replay without initial judgments. Hence, their final judgment cannot be influenced by anything but the play itself.

Since the 2015 season, coaches can challenge most of the plays (fair or foul, in or out of the ballpark, and fan interference), except for pitch calls. This expansion allowed us to quantify most of the effects of discrimination suffered by minority players in discrimination research.

Figure 2 summarizes what types of plays in baseball are reviewable and what are not. The Manager Challenge system enabled them to overturn the judgments, even if the initial ones were biased. There are, however, some exceptions: checked swings[8], pitch calls, and some easy plays[9] are not reviewable. Among them, the number of plays of pitch call is far more significant than other unreviewable ones, so we assume that analyzing the bias of this context reveals almost the full effects of in-group bias in MLB games.

3.4 What is Unique to See This Context?

Figure 3 illustrates the effects of discrimination against minority athletes. Prejudice in pitch calls leads to disadvantages for minorities through two paths: direct and indirect effects.

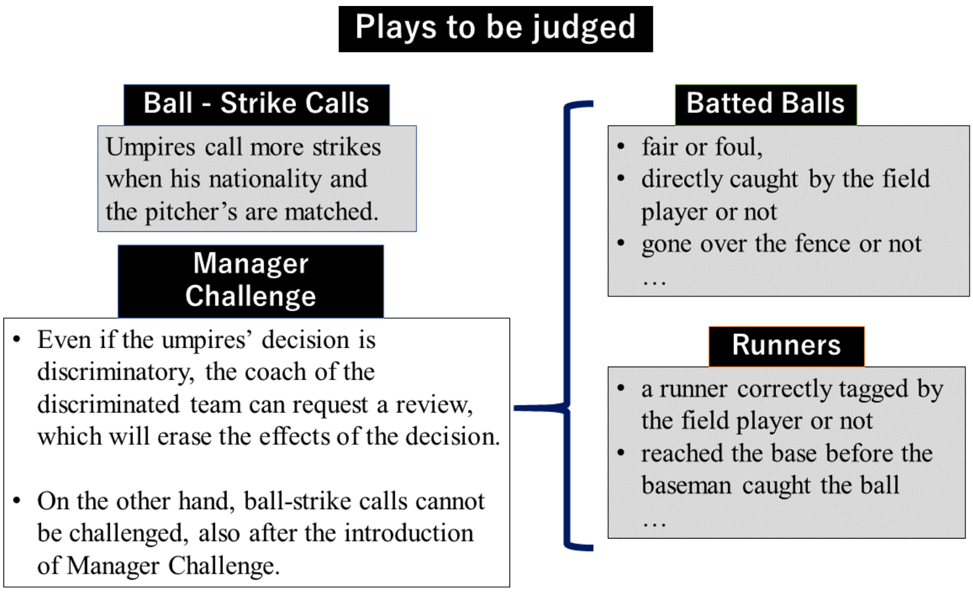
The direct impact means that an umpire's decision conceals a player's ability to perform better at-bat or in the game. As previous sections have mentioned, a strike call instead of a ball has

[7] Some information on fielders' defense is unavailable to the public; however, we can observe aggregate data on television, through Internet coverage or data sites such as Baseball Savant.

[8] Checked swings: When the batter stops swinging midway, umpires judge whether he swang or not.

[9] When the plays occurred in the infield (enough close to the umpires to watch the plays), the rights of the managers to claim reviews are restricted. Section 6 also discusses the magnitude of the impact of these plays.

Figure 2 Illustration of the umpire review



a significantly negative impact on the batter's performance (if it is the third strike, it ends his chance with a strikeout.). It makes it impossible for minority players to perform as well as they would have without such discrimination.

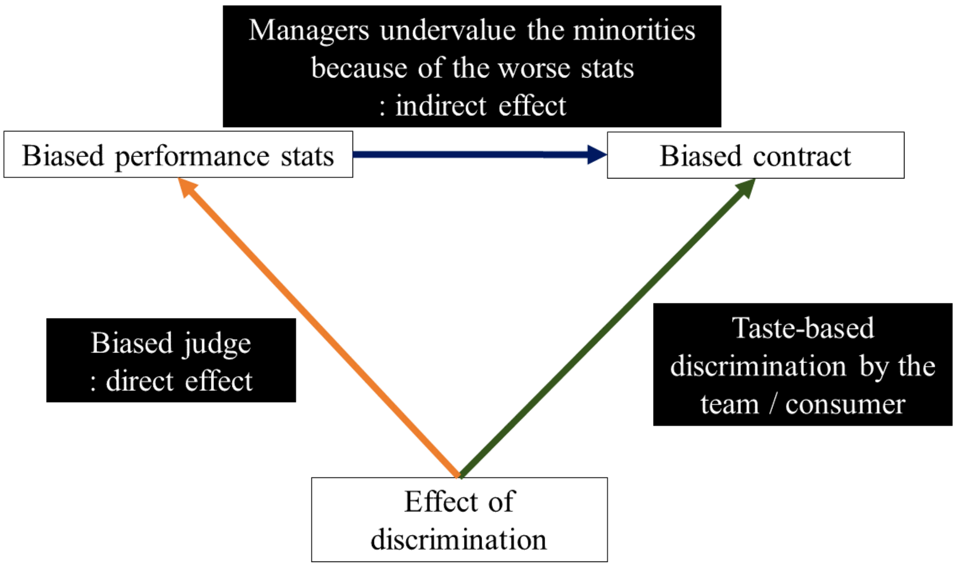
Conversely, the indirect impact corresponds to the arrow extending from "Biased performance statistics" to "Biased contract" in Figure 3. The indirect effect occurs when the team to which the players belong offers (or opts out) a contract for the following year based on performance statistics. Managers evaluate minority players according to their performance statistics; however, these statistics are already compromised as a result of biased umpire calls. Managers are often unaware of the observed downward bias in performance metrics—even if they are aware of it, they do not need to take it into account—and offer lower-value contracts to players. Therefore, even if the team's management does not discriminate against minorities, [10] it indirectly harms players.

This study aims to distinguish between three types of discrimination (direct/indirect effects of biased calls and discrimination in player evaluation by team management) in Figure 3 and discusses which sources to prioritize to improve the situation.

As mentioned previously, almost all discrimination in the playfield occurs in pitch calls. Therefore, analyzing such calls allows us to quantify the conclusive impact of discrimination. We believe this study is the first to address this question, which was also mentioned in Parsons et al. (2011).

[10] Whether they are caused by taste-based discrimination or consumer discrimination is not an issue here.

Figure 3　Path of the Effects of Discrimination against Minority Players



4　Data

This section describes the data used in the analysis. Creating the necessary variables requires tracking (1) the data on each pitch, including pitch location; (2) data on the country of origin of players and umpires; and (3) data on player performance statistics and contracts. We merged them and built a unique dataset.

The analysis will cover the regular-season games in the 2016-2018 seasons. Note that we take a one-year lag about the information on the monetary compensation (See Section 4.3) because general managers design player contracts based on the previous year's performance. For example, the annual salary for 2017–2019 is merged with the statistics for 2016–2018, respectively. [11].

4.1　Pitch Tracking Data

First, we start to explain pitch tracking data. Baseball Savant aggregates tracking data acquired during MLB games and publishes quantitative analysis that predicts players' performance using

[11] In 2020, because of the COVID-19 outbreak, the number of regular-season games changed from 162 to 60. Moreover, the player's earnest money was limited to 37 percent of the contract they originally signed (see How 60-game 2020 MLB season impacts salaries of baseball's highest- and lowest-paid players). Additionally, games of the MiLB, the MLB's subordinate organization, were canceled. Each player was then given the right to decide on participation in the season owing to health risks; several players withdrew from the games themselves. Therefore, the study excluded 2019 and 2020 data from the sample.

physical information such as batted ball velocity and qualitative characteristics of pitches. Much of the data used in the study is available to the public, and we can visit the site to get pitch-by-pitch datasets tied to game and player information.

The total number of games played from 2016 to 2018 was 9,722 and the number of pitches thrown during the games was 20,189,706. In our main analysis, we will focus on the pitches that batters forfeited without swinging at ($N = 1,130,048$).

Table 2 # of pitches with potential bias

Umpire	North America	Other
Umpire, Batter, and Pitcher from the same region	538,117	5,266
Only the umpire from other region	80,260	44,520
Possible bias to the batter	266,357	11,930
Possible bias to the pitcher	163,313	20,285

Table 2 lists the number of pitches devided by plate umpires, pitchers, and the region the batter is from. As discussed below, the majority of MLB certified umpires are from the U.S., so the ratio is somewhat skewed, but in situations where either the pitcher or the batter is from the same country with the place umpire, i.e., where plate umpire discrimination is likely to occur, the number of pitches thrown should be at least 10,000. A sample of about 10,000 pitches can be obtained.

4.2 Nationalities of the Players and Umpires

Table 3 and 4 display summaries of the birthplaces the players and the umpires are from. During the sample period, 2475 players and 108 umpires are rostered as the MLB players and umpires, respectively. The vast majority about both the players and the umpires is the United States, especially for the umpires. Umpires from the United States account for almost 90% of all the umpires.

Chadwick Baseball Bureau contains almost all the players and officials involved in modern American sports. The published dataset includes the year in which a player or an umpire first/last played an MLB game, as well as the unique ID of each player on major websites such as Baseball Savant, Baseball Reference, and Fangraphs. Using the data, the author can merge the tracking data described above with information on the country of origin of players and umpires scraped from Baseball Reference, and data from Fangraphs as described below.

One of the most important determinants of the strike zone is the height of the batter. If the batter's height varies significantly depending on the player's region of origin, these characteristics may be the source of differences based on race. However, we did not identify any statistically

	country	#player	#umpire	#pitches as pitcher	#pitches as batter	#pitches as umpire	Region Group
1	Aruba	1	0	0	4572	0	Latin
2	Australia	6	1	3439	37	0	others
3	Brazil	5	0	552	2911	0	Latin
4	Canada	28	1	9400	12006	12591	North America
5	Colombia	16	0	10973	1882	0	Latin
6	Cuba	48	1	9117	42625	13868	Latin
7	Curacao	5	0	1673	9567	0	Latin
8	Dominican Republic	265	1	113831	124744	9371	Latin
9	Germany	5	0	2245	6649	0	Europe
10	Guam	1	0	350	5	0	others
11	Honduras	1	0	0	0	0	Latin
12	Hong Kong	1	0	665	14	0	Asia
13	Jamaica	1	1	0	0	12374	Latin
14	Japan	16	0	16101	4193	0	Asia
15	Lithuania	1	0	502	0	0	Europe
16	Mexico	24	1	26121	1753	14177	Latin
17	Netherlands	1	0	0	2878	0	Europe
18	Nicaragua	5	0	4225	1545	0	Latin
19	Panama	15	0	4078	3843	0	Latin
20	Peru	1	0	0	0	0	Latin
21	Puerto Rico	43	1	9726	33863	4691	Latin
22	Saudi Arabia	1	0	1534	0	0	others
23	South Africa	2	0	0	219	0	others
24	South Korea	10	0	3696	9222	0	Asia
25	Taiwan	10	0	3165	452	0	Asia
26	U. S. Virgin Islands	3	0	1554	758	0	Latin
27	United Kingdom	2	0	79	5	0	Europe
28	United States	1783	97	878206	762025	1041454	North America
29	Venezuela	175	2	55012	130476	28158	Latin

Table 3 # players and # pitches by birthplace

significant differences in mean height between regions in our sample. This may be due to a survival bias, in which a certain level of height is required for a player to be able to compete in MLB games, but in any case, it explains why there is not a significant difference in the physique of the players that the umpires actually face.

4.3 Contracts

The data for player contracts and the performance are obtained from Baseball Reference and Fangraphs. These are privately owned and operated sites that calculate and publish objective player evaluation indicators based on sabermetrics.

Only a few players are able to play under a season-long contract in MLB, and many are temporarily promoted to MLB and then fired or demoted to MiLB and re-signed. As a result, the number of players for whom we have data on salary during the period in question is small compared to the number of players who have played at least one game. In this section, we will discuss the results of our analysis. The sample size was 611.

Country/Regions	Region Group	#players	Ratio	% pitchers	Height (cm)	s.d.	Weight (cm)	s.d.
United States	North America	1783	72.04	57.88	187.87	(5.71)	94.64	(8.86)
Dominican Republic	Latin	265	10.71	58.49	186.31	(5.34)	94.60	(11.47)
Venezuela	Latin	175	7.07	43.43	184.06	(5.46)	94.14	(11.81)
Cuba	Latin	48	1.94	35.42	185.69	(5.82)	94.15	(9.99)
Puerto Rico	Latin	43	1.74	25.58	184.37	(5.91)	93.02	(11.31)
Canada	North America	28	1.13	57.14	190.25	(5.67)	99.89	(7.89)
Mexico	Latin	24	0.97	83.33	184.46	(4.09)	90.62	(8.90)
Colombia	Latin	16	0.65	50.00	183.31	(6.75)	94.38	(4.40)
Japan	Asia	16	0.65	75.00	184.31	(5.76)	86.50	(6.83)
Panama	Latin	15	0.61	60.00	184.80	(4.93)	90.47	(12.86)
South Korea	Asia	10	0.40	20.00	184.80	(4.52)	100.40	(9.49)
Taiwan	Asia	10	0.40	80.00	184.20	(4.71)	84.90	(11.30)
Australia	others	6	0.24	83.33	186.50	(2.59)	94.50	(9.57)
Brazil	Latin	5	0.20	60.00	188.40	(0.89)	99.40	(12.10)
Curacao	Latin	5	0.20	20.00	185.00	(8.34)	94.00	(17.61)
Germany	Europe	5	0.20	20.00	190.00	(4.42)	101.00	(6.78)
Nicaragua	Latin	5	0.20	60.00	183.40	(7.60)	93.00	(14.40)
U. S. Virgin Islands	Latin	3	0.12	66.67	187.00	(8.19)	96.00	(8.72)
South Africa	others	2	0.08	50.00	181.50	(12.02)	86.50	(4.95)
United Kingdom	Europe	2	0.08	100.00	185.00	(7.07)	94.50	(10.61)
Aruba	Latin	1	0.04	0.00	185.00	(NA)	95.00	(NA)
Guam	others	1	0.04	100.00	190.00	(NA)	99.00	(NA)
Honduras	Latin	1	0.04	0.00	183.00	(NA)	72.00	(NA)
Hong Kong	Asia	1	0.04	100.00	193.00	(NA)	106.00	(NA)
Jamaica	Latin	1	0.04	100.00	198.00	(NA)	117.00	(NA)
Lithuania	Europe	1	0.04	100.00	190.00	(NA)	102.00	(NA)
Netherlands	Europe	1	0.04	0.00	190.00	(NA)	92.00	(NA)
Peru	Latin	1	0.04	100.00	183.00	(NA)	94.00	(NA)
Saudi Arabia	others	1	0.04	100.00	183.00	(NA)	99.00	(NA)

Table 4　Countries Players from

Table 5　Performance Stats for the Pitchers

	All	Asia	Europe	Latin	North America	others
# of Pitchers	1166	19	3	262	877	5
Strike%	63.6	65.5	60.9	63.4	63.6	62.9
Games as a Pitcher	53.06	56.58	35.33	51.23	53.22	118.40
s.d.	(54.33)	(54.34)	(28.11)	(55.38)	(53.83)	(73.79)
Innings Pitched	111.54	146.60	104.00	102.30	113.49	123.93
s.d.	(134.34)	(156.93)	(130.10)	(128.92)	(135.76)	(60.01)
Runs Above Replacement	10.5	18.1	-1.4	9.6	10.7	7.3
s.d.	(23.1)	(28.4)	(5.1)	(20.2)	(23.9)	(12.3)
Wins Above Replacement	1.1	1.9	-0.1	1.0	1.1	0.8
s.d.	(2.4)	(3.0)	(0.6)	(2.1)	(2.5)	(1.3)

5　Identification Strategy

This section describes the empirical strategy to identify the existance and the size of the effect of umpires' discrimination. Using the variables described in Section 4, the estimation equation using the ordinary least squares method is defined as follows.

Table 6 Performance Stats for the Batters

	All	Asia	Europe	Latin	North America	others
# Batters	1957	30	7	494	1419	7
Games Attended	108.73	95.97	192.86	126.58	102.45	91.86
s.d	(119.76)	(85.49)	(176.30)	(131.91)	(115.06)	(75.92)
Plate Appearances	283.60	207.20	678.71	366.21	255.83	15.71
s.d	(503.86)	(332.83)	(759.28)	(563.91)	(480.14)	(30.50)
Batting Runs Created	-1.2	-1.4	-2.1	-1.7	-1.0	-1.3
s.d	(15.7)	(9.3)	(10.2)	(16.1)	(15.7)	(2.2)
Baserunning Runs Created	0.0	-0.3	1.2	-0.5	0.2	0.1
s.d.	(3.3)	(1.4)	(5.5)	(3.9)	(3.1)	(0.1)
Fielding Runs Created	0.0	-0.8	2.1	-0.1	0.1	-0.1
s.d.	(5.9)	(3.1)	(7.1)	(7.0)	(5.6)	(0.3)
Runs Above Replacement	8.6	3.4	25.3	10.4	8.0	-0.6
s.d.	(26.5)	(9.6)	(43.7)	(28.1)	(26.1)	(1.2)
Wins Above Replacement	0.9	0.4	2.5	1.1	0.8	-0.1
s.d	(2.7)	(1.0)	(4.4)	(2.9)	(2.7)	(0.1)

Table 7 Salary and Ages

Region	# players	% Players Salary Available	Mean Salary	s.d.	Mean Age	s.d
Asia	37	56.8	17155880.95	(21305447.82)	33.24	(5.08)
Europe	9	66.7	6232000.00	(9561010.80)	28.67	(3.33)
Latin	608	57.1	9169858.09	(14556144.82)	28.92	(4.51)
North America	1811	52.8	9080632.93	(15573221.13)	29.78	(3.80)
others	10	40.0	2781350.00	(2096420.27)	32.75	(4.99)

For each individual pitch i thrown by pitcher p, to batter b, and called by plate umpire u,

$$\begin{aligned}\mathbf{1}(\text{Strike}_i|\text{Called}_i) = \beta_1 + \beta_2\text{Nationality}_i + \beta_3\hat{f}(\text{plate\_x}_i, \text{plate\_z}_i) \\ + \beta_4\mathbf{X}_i + \delta_b + \gamma_p + \lambda_u + u_i. \qquad (1)\end{aligned}$$

$\mathbf{1}(\text{Strike}_i|\text{Called}_i)$ is a dummy for that the pitch i is called strike, given that the batter takes the pitch. The parameters of interest are β_2, which correspond to the average effects of the nationality of the pitcher, the batter, and the umpire. $\mathbf{X}_i$ stands for the pitch-, game-, or situation-specific characteristics, and δ, γ, and λ are the pitcher, batter, and umpire's fixed effects, respectively.

The nationality dummies include roughly 4 statuses in sum:

1. All of the related agents (the pitcher, the batter, and the plate umpire) are from the same country (baseline),
2. Either the pitcher or the batter is from the same country as the plate umpire while the other is from another one,
3. The pitcher and the batter are from the same country, and the plate umpire is not,
4. Both the pitcher and the batter are from different countries with the plate umpire.

Theoretically, as a whole, plate umpires are likely to favor players from the same country. Therefore, discrimination by the umpire may occur in the situation (2). Each case, furthermore, is classified according to the nationality of the plate umpires. The vast majority of the official MLB umpires are from the United States, so those from other countries are regarded as a minority. This paper distinguishes discrimination against the minority from one by minority umpires against the majority (from the United States) players.

The third term of the equation (1) is the predicted value of the average probability that the pitch is called strike, using the 2-dimensional pitch location. This probability is expressed by $\Pr(\text{StrikeCalled}|\text{plate\_x}, \text{plate\_z})$, where plate_x and plate_z are the horizontal and vertical locations of each pitch. As is described in Section 3.1., It is not unusual for two pitches that pass through the exact same area to be judged differently. Thus, to estimate the effect of the umpire discrimination quantitatively, we should control the average probability of the called strikes given the pitch locations, instead of comparing the ratio of the called strikes among different nationality statuses. To control the information from the pitch location linearly, we use a generalized additive model (GAM) to estimate the density function of strike-call, conditional on the location of the pitches[12]. The following is the equation to estimate the value of $\hat{f}$.

$$1\{\text{StrikeCalled}\} \sim \text{Bin}\left(\theta = \frac{1}{1+\exp(-g(\mathbf{x}))}\right) \tag{2}$$
$$\text{where } g(\mathbf{x}) = s(\text{plate\_x}) + s(\text{plate\_x}, \text{plate\_z}) + s(\text{plate\_z}) \tag{3}$$

The function $g(.)$ in equation (3) is a two-dimensional spline regression of the pitch locations. Taking logit transformation of $g(.)$ (equation (2)), we can obtain the estimated probability of called strikes for each individual pitch with pitch location. Including this term into the OLS regression in equation (1), we can quantify the impact of nationality on pitch calls. If there is no plate-umpire discrimination against the minority players, no deviation is to be observed from the average probability for each location: the rules regarding pitch calls state that plate umpires are to call pitches according only to the pitch locations.

Almost all of the studies on situational strike zone using tracking data suggest that the size of the strike zone varies with pitch count and with the batter's handedness. The latter variation is due to the reversal of the batter's standing position[13]. On the other hand, the former difference is caused by the psychology of umpires, called omission bias. They tend to avoid judging strikes

[12] This methodology is often used to quantify the catchers' technique to get more called strikes.

[13] Figure 4 and 5 suggest that the strike zone is not perfectly symmetrical between right and left handed batters. This is said to be related to the fact that the catcher's catching technique (called "framing") is also likely to affect the pitch call, and that most catchers who play in MLB are right-handed. In any case, this paper assumes that they do not make a critical difference in the same at-bat.

when two strikes and balls when three balls, which are decisions that end the at-bat without the batter hitting the ball into fair territory.

Figure 4 and 5 illustrate the predicted call probability by batter-handedness (The left-hand side for typical left-handed batters, while the right-hand side is for the typical right-handed batters.). To show that pitch counts dramatically affect the average probabilities, we show two anecdotal cases: 0 balls and 2 strikes (0-2 counts) and 3 balls and 0 strikes (3-0). Graphically, you can see that in 0 - 2 counts, umpires favor batters: less likely to call a strike. Figure 5 suggests the opposite implication in 3-0 counts. These primitive results are consistent with previous literature on the umpires in sabermetrics[14].

As will be discussed in Section 6, the estimated probability using GAM has a very large effect on the actual pitch call. In this paper, the sample is divided into 24 subsamples of 12 (0-2 strikes $\times$ 0-3 balls) $\times$ 2, distinguishing between pitch counts and batter left and right, and the probability of each is estimated by GAM and included in the regression in equation (1).

6 Results

6.1 Main Result

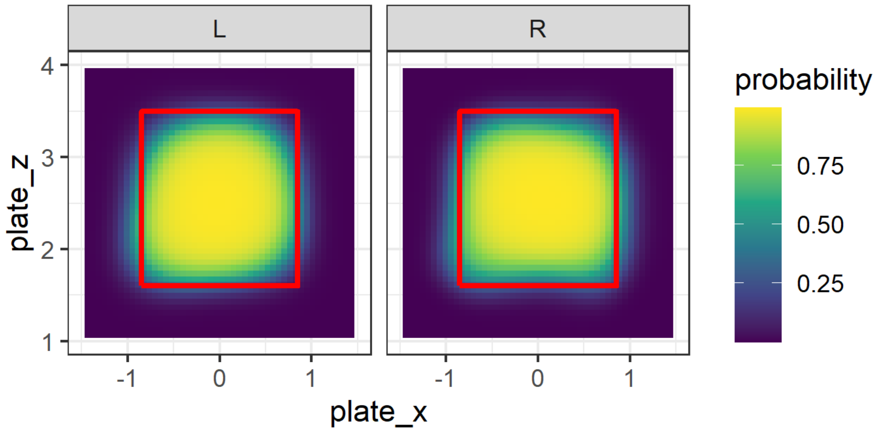
This section discusses the effect of bias due to plate umpire on pitch calls.

Initially, we present results for the sample where the plate umpire is from the United States. Table 8 shows the results of the estimated equation 1. The nationalities of the pitcher and batter are listed in rows 3 through 5, with the texts before the underscore (_) indicating the nationality of the pitcher, and the nationality of the batter follows after it. For example, "NoUS_US" indicates that the batter is from the United States and the pitcher is from elsewhere. The pitcher's choice and the manager's decision on which pitcher to play against which batter are endogenous variables. Including these variables may be "bad controls," as pointed out by Angrist and Pischke (2008). In the first column of Table 8, we show the results with only the estimated strike call probability $\hat{f}$ and nationality information as explanatory variables.

The estimation results were consistent with the hypothesis. When the pitcher is from the U.S. and the batter is from elsewhere, the plate umpire from the United States tends to decide in favor of the pitcher. Specifically, an average deviation of 0.3-0.4 percentage points from the probability of a strike call estimated solely by the location of the pitch's passage occurs, i.e., the batter is more likely to have a strike called. The estimates were statistically significant (at

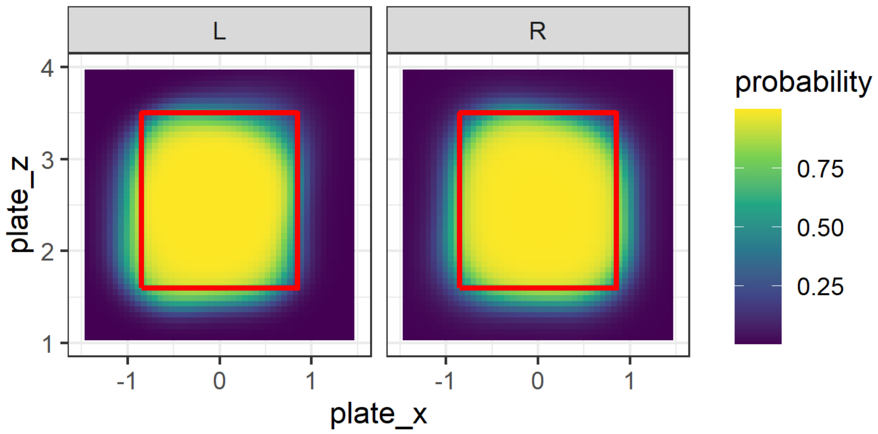
[14] The term "sabermetrics" refers to an attempt to analyze baseball from an objective standpoint using scientific methods and to gain new insights that could not be obtained with the old methodology of analysis based solely on experience and feeling.

Figure 4 Predicted Strike-Call Probability: 0 Balls and 2 Strikes



The probability of a strike call when the pitch count is 0-2,
i.e., the pitcher has an advantage.

Figure 5 Predicted Strike-Call Probability: 3 Balls and 0 Strikes



The probability of a strike call when the pitch count is 3-0,
i.e., the batter has an advantage.
You see that it is more likely that pitches on the edge
of the strike zone is called strike, compared to when 0-2 counts.

the 0.1% level). On the other hand, if the pitcher is from a non-U.S. country and the batter is from the United States, the pitch call will favor the batter. However, the point estimates were relatively small for batters, and the estimation results in column 4, which contained the most explanatory variables, were no longer statistically significant.

The result where both the pitcher and the batter are not from the United States does not support

Table 8 Discrimination by the American Umpires

	(1)	(2)	(3)	(4)	(5)
(Intercept)	0.002 ***	0.002 ***			
	(0.000)	(0.000)			
prob	0.992 ***	0.992 ***	0.992 ***	0.991 ***	0.991 ***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
NoUS_NoUS	0.001	0.001	0.002 *	0.002 *	0.003 **
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
US_NoUS	0.003 ***	0.003 ***	0.004 ***	0.004 ***	0.004 ***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
NoUS_US	-0.002 *	-0.002 *	-0.002 *	-0.001 *	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
pitcher_hitting			0.021 ***	0.022 ***	0.022 ***
			(0.002)	(0.002)	(0.002)
away_batting				0.004 ***	0.004 ***
				(0.000)	(0.000)
Fixed Effects			catcher	catcher	catcher, pitch_type
#observations	1048047	1048047	1048047	1043840	1043840
R squared	0.751	0.751	0.752	0.752	0.752
F statistic	1021865.000	1021865.000			
P value	0.000	0.000			

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

the hypothesis. In this matchup with no intuitive conflict with the umpire the theory predicts to show similar results to the baseline case where both pitcher and batter were from the United States. However, some of the results suggested that the plate umpire would call in favor of the pitcher.

In addition, the results pointed to count-dependent changes in the strike zone and the existence of a hometown bias. These are consistent with the previous results. For example, the percentage of strike calls increases by 0.4 percentage points in the innings table, where most games hit the visiting team. Although these are not the main focus of this study, the fact that they are consistent with previous studies implies the validity of the analysis in this study. In particular, the effect of pitch count is much larger than the ones of player nationality and other factors, with approximately tenfold variation occurring.

As shown in Figures 4 and 5, the probability of an average strike call changes rapidly at the edge of the strike zone. The plate umpire is also less sensitive to nationality since a pitch that passes through a position with an extremely high or low average probability is a strike ball for all. Table 9 shows the estimation results when the sample is restricted to pitches with average strike call probabilities $\hat{f}$ between 30% and 70%. This region is called the "shadow zone" and is empirically considered to be the pitches for which the strike zone is likely to change depending on the context.

Table 9 Edge of the Strike Zone

	(1)	(2)	(3)	(4)	(5)
(Intercept)	0.011	0.011			
	(0.007)	(0.007)			
prob	0.919 ***	0.919 ***	0.918 ***	0.919 ***	0.919 ***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
NoUS_NoUS	0.007	0.007	0.010	0.009	0.012 *
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
US_NoUS	0.008 *	0.008 *	0.011 **	0.011 **	0.012 **
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
NoUS_US	-0.005	-0.005	-0.004	-0.005	-0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
pitcher_hitting			0.069 ***	0.074 ***	0.067 ***
			(0.009)	(0.010)	(0.010)
away_batting				0.014 ***	0.014 ***
				(0.003)	(0.003)
Fixed Effects			catcher	catcher	catcher, pitch_type
#observations	100359	100359	100359	99994	99994
R squared	0.054	0.054	0.064	0.065	0.070
F statistic	1218.814	1218.814			
P value	0.000	0.000			

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

The results are consistent with our prediction, especially in situations where batters are discriminated against, and the point estimates show 0.8 percentage points to 1.2 percentage points. In other words, the estimates in the full sample are relatively large due to this part of the effect. However, the effect of discrimination was also present for obvious strikes and obvious balls are thrown with an average strike call probability of less than 10 % or more than 90 %.

The last result of the main estimation is about the judgement of the non-American plate umpire. Table 10 shows the OLS results with the sample restricted to umpires from countries other than the US. Statistically, the results do not provide strong support for reverse discrimination by minorities. None of the dummy variables for nationality were dominant at the 10% level. However, the signs of the point estimates are all consistent with the hypotheses, especially when the pitcher is non-American and the batter is American, and the coefficient estimates are relatively large, ranging from 0.3 percentage points to 0.6 percentage points.

These results indicate that plate umpires from the U.S. make pitch calls that are influenced by nationalistic bias based on the country of origin of the pitcher and batter. The effect is particularly large in situations where batters are discriminated against, and they are more likely to underperform themselves by being called more strikes.

Table 10 Reverse Discrimination

	(1)	(2)	(3)	(4)	(5)
(Intercept)	-0.002	-0.002			
	(0.003)	(0.003)			
prob	0.984 ***	0.984 ***	0.984 ***	0.983 ***	0.984 ***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
NoUS_NoUS	0.003	0.003	0.001	0.001	-0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
US_NoUS	-0.000	-0.000	-0.001	-0.001	-0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
NoUS_US	0.006	0.006	0.005	0.005	0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
pitcher_hitting			0.025 ***	0.027 ***	0.025 ***
			(0.006)	(0.006)	(0.006)
away_batting				0.004 *	0.003 *
				(0.002)	(0.002)
Fixed Effects			catcher	catcher	catcher, pitch_type
#observations	82001	82001	82001	81796	81796
R squared	0.743	0.743	0.745	0.745	0.745
F statistic	72288.430	72288.430			
P value	0.000	0.000			

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

6.2 Size of Effect: Indirect Effects on Players' Salary

Next, we discuss the average impact of discrimination in size: from Section 6.1. when the plate umpire and pitcher are from the U.S. and the batter is from the rest of the world, the probability of a strike call increases by .4 percentage points per pitch. In this paper, we quantify how much of a financial loss this causes to minority players.

Table 11 quantifies the impact of improvements in player performance metrics on subsequent seasons using data on hitter performance from 2016 to 2018 and annual salary the following year. Specifically, we regress the following equation.

$$W_i = \alpha + \beta \text{Nationality}_i + \gamma \text{PerfObs}_i + \mathbf{bX}_i \tag{4}$$

To account for issues such as multi-year contracts and applications for annual salary arbitration, we sum up performance measures and the annual salary data over the three years. W_i in equation (4) represents the annual salary over the three years, and PerfObs_i is a proxy variable for the performance measure used by team managers to consider the annual salary for their players. In

Table 11 The Return to Performance on 3-year Conpensation

	salary	log-salary	salary	log-salary
Age	5207970.404 **	0.621 ***	6876505.684 ***	0.813 ***
	(1827345.814)	(0.154)	(1616971.340)	(0.151)
Age_squared	-58552.821	-0.007 **	-88374.735 **	-0.011 ***
	(32027.095)	(0.003)	(27932.541)	(0.003)
Asia	-953576.469	0.264	-3956462.722	0.012
	(6824475.892)	(0.345)	(6526262.934)	(0.345)
Latin	2162256.495	0.342 ***	3186830.082 **	0.442 ***
	(1157596.818)	(0.095)	(1181647.678)	(0.117)
Europe	-1614844.584	0.226	2479869.411	0.613
	(1222698.060)	(0.513)	(4167365.210)	(0.747)
RAR	224003.728 ***	0.024 ***		
	(18381.413)	(0.001)		
Bat			276739.262 ***	0.023 ***
			(35286.391)	(0.002)
BsR			-71511.698	0.009
			(129943.009)	(0.010)
Fld			264458.450 ***	0.026 ***
			(66378.611)	(0.005)
Pos			-137458.304	-0.002
			(80851.681)	(0.006)
#observations	611	611	611	611
R squared	0.437	0.570	0.398	0.378
F statistic	63.089	183.137	31.027	57.595
P value	0.000	0.000	0.000	0.000

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

other words, this value is subject to umpire bias.

PerfObs_i can be further decomposed into the following definition.

$$\text{PerfObs}_i = \text{PerfPot}_i + \phi \times \text{Attendance}_i \tag{5}$$

PerfPot_i represents the potential, i.e., the true ability of the batter, that would be observed if there were no discrimination by the umpire. Attendance_i represents the number of opportunities distributed to the player. It can be interpreted, in other words, as the frequency with which players who would be subject to discrimination appear on the field. The value of Φ represents the impact of discrimination as estimated in the main results section and indicates that the more minority batters are plate-appearances, the greater the downward bias on the measure of run creation for that player.

The two performance indicators used in this paper are "Runs Above Replacement (RAR)" and "Bat (Batting)." RAR is a measure of "how many more points a player contributes than a replacement-level[15] player", and the number can be converted to points per score (runs) in baseball. Therefore, the coefficient value of RAR in equation (4) represents the rate of increase in annual salary when the player's contribution to the team's runs increases by one point (over three years). In short, it stands for the return on contribution. "Bat" is a similar measure of a player's hitting-related ability on a scale of scoring contributions created.

As shown in Equation (4), this regression also includes a proxy variable for the player's home region. This corresponds to the effect of the player's region of origin on the annual salary amount, which remains even after accounting for differences in the player's track record. The estimated coefficients of these dummy variables are statistically significant if there is an effect of discrimination against minority players by team management or consumers (see Figure 3).

Consider the results in Table 11. First, there was no strong support for the possibility of discrimination by team or fan, at least when controlling for factors such as player performance and age. None of the estimates for the regional dummies were statistically significant, or represented treatment that favored minority players. The remainder of the discussion will focus on the effects of discrimination by ballplayers. Rows 5 and 6 of Table 11 show the return of runs scored by players relative to their annual salary. It shows that the return on an increase of one point in runs scored by a player over three years is estimated to be about $224,000, or 2.4%. The coefficient values are statistically significant (at the 0.1% level) when using either RAR or Bat, and the point estimates show similar impact size. As mentioned earlier, since these indicators implies a player's ability in the form of runs, quantifying the negative runs creations produced by pitch-calling bias would allow us to measure the impact of the decision in monetary value.

The impact of pitch calls on the creation of runs is done using the concept of Runs Expectancy[16] used in MLB. We can regard baseball as a dynamic game of discrete time in which a pitcher and a batter play against each other with intervals in between. For example, a double by the batter at the beginning of an inning changes the situation from no outs and no runners on base to no outs and no runners on second base. For each state[17], if we average the runs scored going into the rest of the inning, we get the Runs Expectancy table as shown in Table 12. For example, with bases loaded, if there are no outs, we can expect to score 2.23 more runs on average in that inning. However, if there is one out, that number decreases to 1.55 runs.In other words, striking out with the bases loaded and no outs create a negative score of 0.68 (1.55 - 2.23). By

[15] The replacement level indicates the level of competence of the players in the MLB player market who are being sought, i.e., those whom teams can be immediately acquire by paying minimum guaranteed salary.It will not be explained in detail here. It is sufficient to understand that a player's value can be evaluated in terms of points.

[16] For more information on the concept of expected runs scored, please refer to the Sabermetrics website.

[17] Out count: 0, 1, or 2 $\times$ 3 with or without runners on base, 8 ways = 24 different states defined.

averaging this change in expected score for example, strikeouts, we can estimate the average (state-independent) run creation value of a strikeout. Similarly, we can calculate how many points the expected score would be reduced (A strike is a decision that disadvantages the batter.) on average if the umpire calls a strike instead of a ball when the batter takes the ball. According to this, the average scoring creation for being taken one strike is about -1.4 points.

Table 12 Runs Expectancy by 24 States

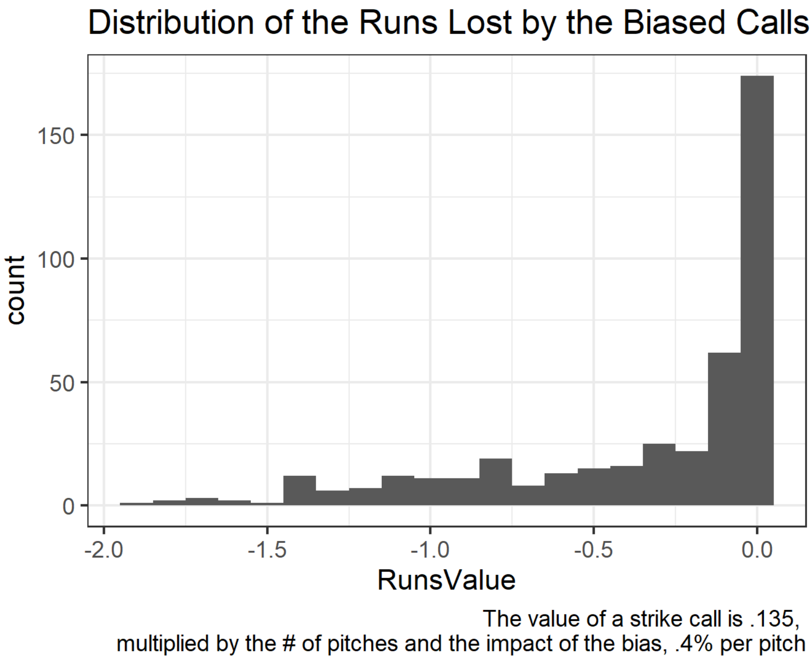
Outs/Runners_on	0	1	2
___	0.50	0.27	0.10
__3	1.39	0.97	0.37
_2_	1.14	0.69	0.32
_23	1.98	1.40	0.56
1__	0.88	0.52	0.22
1_3	1.78	1.21	0.49
12_	1.46	0.93	0.43
123	2.23	1.55	0.76

Minority batters are on average 0.4 percentage points more likely to have a strike called per pitch, so for example, 100 pitches against an American-born player with an American-born umpire would score -1.4 points $\times$ 0.004 percent per pitch $\times$ 100 pitches = 0.56 points. It will harm the observed ability to runs creation. Figure 6 quantifies the sum of the runs created lost over three years by minority players for whom Salary data are available. Since players with fewer opportunities to bat are also less likely to be discriminated against, many players appear to have little impact on the total over the three years. On the other hand, hitters who are the core players of their teams and bat all year long have more opportunities to bat in situations where they are discriminated against, which means that they lose nearly two runs in three years due to discriminatory pitch calls. In particular, the best players in the top 25 percentile among the minorities lost an average of 0.58 points worth of runs. As shown in Table 11, the return per runs created is $224,000, so the estimated salary they lose over three years is about $130,000. Some of the players signed before free agency sign the contract with about the minimum compensation set at about $570,000. This is a loss that we should not ignore.

6.3 Endogeneity: Bahavioral Response of the Batters

The impact size discussed in the two sections is a lower bound. This is because, in the context of the strike zone, significant endogeneity is noted.

Figure 6 Distribution of the Negative Contribution by the Biased Calls



If minority hitters are empirically (or quantitatively) aware of the existence of discrimination by plate umpires, they will decide whether or not to swing at a pitch given a wider strike zone.

$$U_s = E[\text{Runsvalue}|\text{Contact}] \times Pr(\text{Contact}) + E[\text{Runsvalue}|\text{Missed}] \times (1 - Pr(\text{Contact})) \tag{6}$$

$$\begin{aligned} U_t = & E[\text{Runsvalue}|\text{Called Strike}] \times Pr(\text{Called Strike}) \\ & + E[\text{Runsvalue}|\text{Called Ball}] \times (1 - Pr(\text{Called Strike})) \end{aligned} \tag{7}$$

The batter's decision to swing or take can be described by the following utility function The expressions (6) and (7) represent the incremental scoring contribution created when the batter swings at the pitch, respectively. When a swing is made, if the batter contacts the ball, the expected value of the contribution created by the generated pitch is generated, while a loss of one strike always occurs if the batter swings and misses. On the other hand, if the pitch is taken, it is determined whether the pitch is a ball or a strike based on the average strike call probability $\hat{f}$ estimated by 2. Taking convex combination of the value of one strike and one ball by the average probability of a strike for the pitch location of that pitch, we can calculate the expected value of taking a pitch as measured by scoring creation.

Therefore, the above two equations can be rewritten as a function of pitch location $(\text{plate\_}x, \text{plate\_}z)$ as follows.

$$
\begin{aligned}
U_s =& S(\text{plate\_}x, \text{plate\_}z) && (8) \\
U_t =& E[\text{Runsvalue}|\text{Called Strike}] \times \hat{f}(\text{plate\_}x, \text{plate\_}z) \\
& + E[\text{Runsvalue}|\text{Called Ball}] \times (1 - \hat{f}(\text{plate\_}x, \text{plate\_}z)) && (9)
\end{aligned}
$$

Table 13 shows the variant of the average expected score when a pitch was a strike, a ball, or a foul in each pitch count. Substituting these values into the equation (9) calculates the average value when a pitch is taken.

For example, consider the case of taking a pitch from 0 balls and 2 strikes (e.g. "3-2" represents 3 balls and 2 strikes) such that $\hat{f} = 50\%$. If the pitch is ruled a strike, the value is -0.164; if it is a ball, it is 0.043. Thus, the value of taking this pitch is

$$-0.164 \times 0.5 + 0.043 \times (1 - 0.5) = -0.0605.$$

Table 13 Δ Run Expectation at Each Pitch Count

Count	if strike	if ball	if foul
0-0	-0.042	0.037	-0.042
1-0	-0.055	0.060	-0.055
2-0	-0.079	0.097	-0.079
3-0	-0.083	0.116	-0.083
0-1	-0.059	0.024	-0.059
1-1	-0.040	0.036	-0.040
2-1	-0.042	0.093	-0.042
3-1	-0.003	0.200	-0.003
0-2	-0.164	0.043	0
1-2	-0.207	0.034	0
2-2	-0.241	0.132	0
3-2	-0.373	0.203	0

In a similar way to finding $\hat{f}$, we use the expected value of run creation when a batter swings at a pitch that has passed a specific pitch location, the function $S(.,.)$ can be calculated.

$$S(\text{plate\_x}, \text{plate\_z}) = r(\text{plate\_x}) + r(\text{plate\_x}, \text{plate\_z}) + r(\text{plate\_z}) \tag{10}$$

$r(.,.)$ is the spline regression.

Table 14 Run Values of Each Play

Events	Run Value
Homerun	1.38
Triple	1.06
Dounle	0.76
Error	0.49
Single	0.45
Catcher Interference	0.34
Hit By Pitch	0.33
Bases-on-Balls	0.31
Intentional Walk	0.17
Fielders Choice	-0.23
Strikeout	-0.27
Batted-Ball Outs	-0.27

If the batter strikes out on a pitch, the value of "if strike" in the Table 13 is used as the explained variable according to the pitch count. On the other hand, if the batter hits a double, for example, the average scoring value of a double, 0.76, is the value of the swing. Table 14 shows the average value that each play has.

Using these values, we can quantify the extent to which incentives exist to swing at a ball pitched at a particular pitch location. For a pitch i,

$$\begin{aligned}\text{Incentive}_i =& S(\text{plate\_x}, \text{plate\_z}) - \hat{f}(\text{plate\_}x, \text{plate\_}z) \times \text{Value when Strike}_i \\ & - (1 - \hat{f}(\text{plate\_}x, \text{plate\_}z)) \times \text{Value when Ball}_i \end{aligned} \tag{11}$$

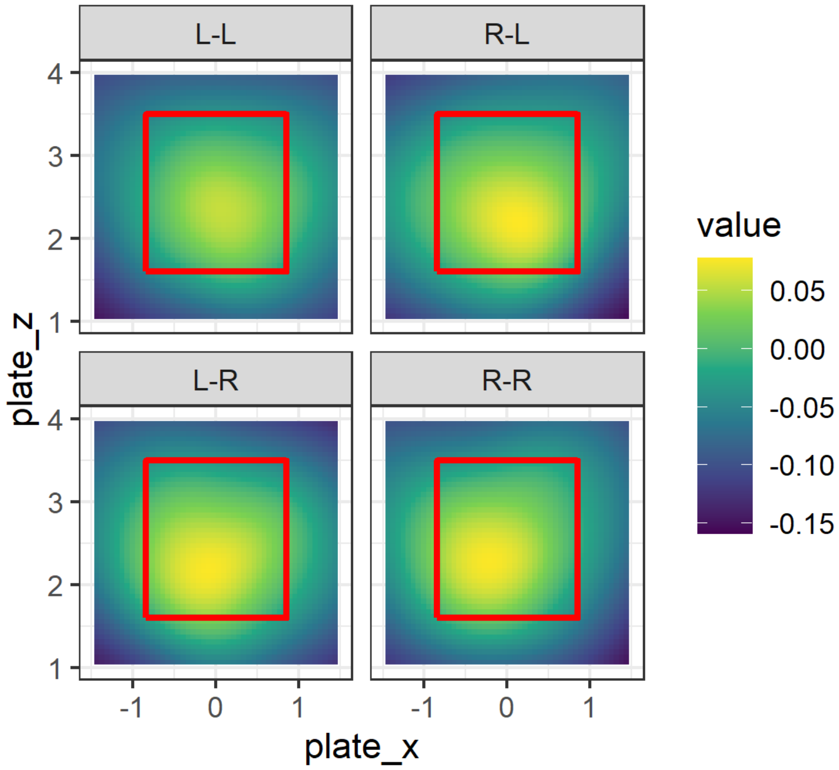
This corresponds to the difference between the equations (8) and (9), where the larger the number, the larger the expected value of runs created by swinging at a pitch rather than taking it. Figure 7 illustrates these values given pitch locations, seperating the handedness of the pitchers and the batters. There are differences depending on the pitcher's or batter's handedness, but in general, you will find that the expected value is higher when you swing around the heart of the strike zone.

Using this, we estimate the following regression.

$$\mathbf{1}(\text{Swung})_i = \beta_1 + \psi \text{Nationality}_i + \beta_2 \text{Incentive}_i + \beta_4 \mathbf{X} \tag{12}$$

Nationality_i corresponds to the nationality of the player and plate umpire, as in equation (1).

Figure 7 Expected Run Values when Swinging



The dependent variable is the dummy that indicates if the batter swings or not. If minority players are aware of the existence of discrimination and expect their probability of calling a strike to increase when facing an American pitcher, they will swing more even in situations where the incentive to swing is relatively small.

Table 15 shows the estimation results for OLS, where columns 1 and 2 include all pitches in the sample, and columns 3 and 4 restrict the sample to the "shadow zone" where the percentage of strike calls predicted by the location of the pitch's path is between 30% and 70%. Since the sample is limited to cases where a minority batter is at bat, there are three nationality states: the pitcher is from the US, the plate umpire is from the US, and both of them are from the US. The estimation results did not support the existence of a change in batter behavior according to nationality. The value of the point estimate varied above and below zero depending on the sample selected, and when all pitches were included in the sample, it even resulted in a rather low percentage of swings in situations where he himself was discriminated against.

7 Discussion

The probability variation of strike calls by nationality bias was generally as predicted by the theory. Many studies dealing with the strike zone context, including Parsons et al. (2011),

Table 15 Behavioral Responses of the Minority Batters

	(1)	(2)	Edge	Edge
US Umpire	-0.003	-0.003	0.017	0.022
	(0.004)	(0.004)	(0.012)	(0.012)
US Pitcher	-0.007	-0.004	0.022	0.034 **
	(0.004)	(0.004)	(0.013)	(0.013)
US Umpire & US Pitcher	-0.009 **	-0.006	0.007	0.020
	(0.004)	(0.004)	(0.012)	(0.012)
Fixed Effects		pitch_type		pitch_type
#observations	739969	737581	80508	79946
R squared	0.247	0.256	0.130	0.149
F statistic	63467.440		2702.338	
P value	0.000		0.000	

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

have used the race of the player as a proxy variable instead of nationality. On the other hand, for example, Sandberg (2017) pointed out the existence of in-group bias by the nationality of the umpire using a dressage dataset. Since some studies in the context of the strike zone use country-of-origin-dependent Latino/Asian identification and lists of black athletes in major media articles as a way to identify their race, whether the bias is due to nationality or race There is room for debate as to whether the bias is due to nationality or race. On the other hand, in MLB games, the names of the players at bat and the pitchers on the mound are called out by speakers in the stadium, and their names are written on the scoreboard attached to the batter's eye. In this situation, the umpire can know their names at any time. Using data from their experiments, Bertrand and Mullainathan (2004) and Edelman et al. (2017) have pointed out that people's names may be used as cues to determine their race. In addition, the plate umpire, in particular, is located at a very close distance to the player and can easily know the skin color of the player. Monk (2015)and others have argued that people's skin color predicts the degree of discrimination they are subjected to, and intuitively, these effects are likely to occur in baseball. Since the faces of athletes are available to the public, further research in this area using such data may provide insights into discriminating between nationality bias and racial discrimination.

8 Conclusion

This paper has used the context of strike calls in the MLB to show evidence for minority discrimination by evaluators (umpires). The in-group bias that occurs when evaluating the achievements of workers (players) also causes workers to suffer economic losses through subjective evaluations by employers (team managers) who enter into contracts with workers based on the bias. This paper suggests that even though umpires are highly skilled professionals, they are affected by their in-group bias. Such a problem can occur even in a more general work environment that lacks objective measures of worker performance and so we can consider our results as a contribution to the labor economics literature using empirical data.

Another advantage of research using MLB data is that it allows us to track the behavior of workers discriminated against, as discussed in Section 6.3. According to the results, batters discriminated against would recognize it and make behavioral changes, while the financial losses caused by discrimination were very severe. The fact that the effects of such bias exist in the MLB labor market, which attracts players from all over the world, may shed some light on the debate over the introduction of non-human entities such as robot umpires.

References

A. Agan and S. Starr. Ban the box, criminal records, and racial discrimination: A field experiment. *Quarterly Journal of Economics*, 133(January):191–235, 2018. doi: 10.1093/qje/qjx028. URL https://doi.org/10.1093/qje/qjx028.

J. D. Angrist and J.-S. Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Dec. 2008. ISBN 0691120358.

J. D. Angrist and J.-S. Pischke. *Mastering' metrics: The path from cause to effect*. Princeton: Princeton University Press, 2014.

J. Archsmith, B. Sampat, J. Archsmith, and M. Neidell. The Dynamics of Inattention in the (Baseball) Field. *Discussion Paper in IZA*, (14440), 2021. URL https://www.iza.org/publications/dp/14440/the-dynamics-of-inattention-in-the-baseball-field.

O. Ashenfelter, R. Oaxaca, and R. Oaxaca. The Economics of Discrimination: Economists Enter the Courtroom. *The American Economic Review*, 77(2):321–325, 1987. ISSN 0002-8282. doi: 10.2307/1805472.

Baseball Reference. URL https://www.baseball-reference.com/.

Baseball Savant. URL https://baseballsavant.mlb.com/.

M. Bertrand and S. Mullainathan. Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *American Economic Review*, 94(4):991–1013, September 2004. doi: 10.1257/0002828042002561. URL https://www.aeaweb.org/articles?id=10.1257/0002828042002561.

Chadwick Baseball Bureau. URL http://chadwick-bureau.com/.

D. L. Chen, T. J. Moskowitz, and K. Shue. Decision Making Under the Gambler' s Fallacy: Evidence from Asylum Judges, Loan Officers, and Baseball Umpires *. *The Quarterly Journal of Economics*, 131(3):1181–1242, 03 2016. ISSN 0033-5533. doi: 10.1093/qje/qjw017. URL https://doi.org/10.1093/qje/qjw017.

B. Edelman, M. Luca, and D. Svirsky. Racial discrimination in the sharing economy: Evidence from a field experiment. *American Economic Journal: Applied Economics*, 9(2):1–22, April 2017. doi: 10.1257/app.20160213. URL https://www.aeaweb.org/articles?id=10.1257/app.20160213.

Fangraphs. URL https://www.fangraphs.com/.

J. H. Fiva and A. H. Halse. Local favoritism in at-large proportional representation systems. *Journal of Public Economics*, 143:15–26, nov 2016. ISSN 0047-2727. doi: 10.1016/J.JPUBECO.2016.08.002.

J. R. Hill, J. Madura, and R. A. Zuber. The short run demand for major league baseball. *Atlantic Economic Journal*, 10(2):31–35, 1982. ISSN 01974254. doi: 10.1007/BF02300065.

P. Holmes. New evidence of salary discrimination in major league baseball. *Labour Economics*, 18(3):320–331, jun 2011. ISSN 09275371. doi: 10.1016/j.labeco.2010.11.009.

How 60-game 2020 MLB season impacts salaries of baseball's highest- and lowest-paid players . URL https://www.cbssports.com/mlb/news/how-60-game-2020-mlb-season-impacts-salaries-of-baseballs-highest-and-lowest-paic

F. Jiang, Y. Qian, and S. E. Yonker. Hometown biased acquisitions. *Journal of Financial and Quantitative Analysis*, 54(5):2017–2051, 2019a. doi: 10.1017/S002210901800131X.

F. Jiang, Y. Qian, and S. E. Yonker. Hometown biased acquisitions. *Journal of Financial and Quantitative Analysis*, 54(5):2017–2051, 2019b. doi: 10.1017/S002210901800131X.

J. W. Kim and B. G. King. Seeing stars: Matthew effects and status bias in major league baseball umpiring. *Management Science*, 60(11):2619–2644, 2014. ISSN 15265501. doi: 10.1287/mnsc.2014.1967.

A. C. Krautmann, E. Gustafson, and L. Hadley. Who pays for minor league training costs? *Contemporary Economic Policy*, 18(1):37–47, 2000. ISSN 10743529. doi: 10.1111/j.1465-7287.2000.tb00004.x.

R. E. McCormick and R. D. Tollison. Why do black basketball players work more for less money? *Journal of Economic Behavior & Organization*, 44(2):201–219, 2001. ISSN

0167-2681. doi: https://doi.org/10.1016/S0167-2681(00)00156-6. URL https://www.sciencedirect.com/science/article/pii/S0167268100001566.

MLB. URL https://www.mlb.com/official-information/about-mlb.

MLB.com: Strike Zone. URL https://www.mlb.com/glossary/rules/strike-zone.

E. P. Monk. The cost of color: Skin color, discrimination, and health among African-Americans. *American Journal of Sociology*, 121(2):396–444, 2015. ISSN 00029602. doi: 10.1086/682162.

Much required to become MLB umpire. URL http://mlb.mlb.com/content/printer_friendly/mlb/y2007/m08/d27/c2173765.jsp.

C. Nardinelli and C. Simon. Customer racial discrimination in the market for memorabilia: The case of baseball. *Quarterly Journal of Economics*, 105(3), 1990. ISSN 15314650. doi: 10.2307/2937891.

C. A. Parsons, J. Sulaeman, M. C. Yates, and D. S. Hamermesh. Strike three: Discrimination, incentives, and evaluation. *American Economic Review*, 101(4):1410–35, June 2011. doi: 10.1257/aer.101.4.1410. URL https://www.aeaweb.org/articles?id=10.1257/aer.101.4.1410.

Retrosheets. URL https://www.retrosheet.org/.

A. Sandberg. Competing Identities: A Field Study of In - group Bias Among Professional Evaluators. *The Economic Journal*, 128(613):2131–2159, 11 2017. ISSN 0013-0133. doi: 10.1111/ecoj.12513. URL https://doi.org/10.1111/ecoj.12513.

G. W. Scully. Pay and performance in major league baseball. *The American Economic Review*, 64(6):915–930, 1974. ISSN 00028282. URL http://www.jstor.org/stable/1815242.

Statcast. URL https://www.mlb.com/glossary/statcast.

S. Tainsky, B. M. Mills, and J. A. Winfree. Further examination of potential discrimination among mlb umpires. *Journal of Sports Economics*, 16(4):353–374, 2015.