Tohoku University Research Center for Policy Design Discussion Paper

TUPD-2025-002

Overrewarding Luck and Lucky Streak: Evidence from MLB

Koji Yashiki
Graduate School of Economics and Management, Tohoku University

February 2025

TUPD Discussion Papers can be downloaded from:

https://www2.econ.tohoku.ac.jp/~PDesign/dp.html

Overrewarding Luck and Lucky Streak: Evidence from MLB *

Koji Yashiki†

Abstract

This study investigates the outcome bias by exploiting natural experimental situations, where outcomes are likely to be exogenously determined, from Major League Baseball. Specifically, I extract hits near the wall where a slight distance affects the outcome and test the difference in evaluation between random successes and failures. The estimation results provide evidence of the persistent outcome bias and its heterogeneity based on player race and manager experience. Managers are more likely to continually allow players who happen to be successful to play in subsequent games. Additionally, they tend to over-evaluate the random success of white players compared to racial minorities, although their decisions become less biased with increased experience. The racial disparities in evaluation may further complicate the problems posed by the outcome bias; however, the improvement in decision-making with experience suggests the potential solution.

JEL Classification: D86; D90; Z20

Keywords: outcome bias; principal agent problem

*I thank Akira Hibiki, Michio Suzuki, and Yoshiyuki Nakazono for their comments and suggestions.

†Tohoku University, yashiki.koji.t6@dc.tohoku.ac.jp

1 Introduction

Does luck affect the manager's evaluation of the worker? Contract theory assumes that the outcome of the worker's performance is equal to effort plus noise (Bolton and Dewatripont, 2005). Typically, the noise is interpreted as luck under an uncertain situation, then disentangling effort and luck is critical when evaluating outcomes (Brownback and Kuhn, 2019). However, in practice, individuals tend to overreact to outcomes without fully distinguishing between effort and luck. That is, the lucky workers who happen to be successful are likely to be evaluated as making more effort, and vice versa. This type of cognitive bias is referred to as the outcome bias (Baron and Hershey, 1988).

This study investigates the outcome bias by extracting a natural experimental setting from professional sports data. The natural experimental setting is the situation where the hits fall near the wall in baseball, which is obtained from high-frequency tracking data in Major League Baseball: MLB. It is reasonable to assume that outcomes of hits near the wall are randomly determined, based on the disciplinary characteristics of baseball and evidence from descriptive statistics. The special situation allows this study to test whether the outcome with no information about the player's effort affects the manager's evaluation of the player. In addition, the large sample enables this study to examine its persistence and heterogeneity by race and experience.

The estimation results show that random success makes more the manager's evaluation of the player. Players whose hits near the wall are homeruns are more likely to play the next game as a starter than those whose hits are outs. Moreover, the lucky players are entrusted with a more important role in the batting order. The effect of over-evaluating random success decays over time, but persists for at least 10 games. Furthermore, some additional analysis shows two heterogeneities in the outcome bias. The first is heterogeneity by player's race, with the outcome bias being larger and more persistent for white players than racial minorities. The second is heterogeneity by manager's experience, with scale and persistence of the outcome bias being smaller for more experienced managers.

Previously, the outcome bias has mainly been studied in laboratory experiments (Baron and Hershey, 1988; Alicke et al., 1994; Alicke and Davis, 2004), and there has been further research into heterogeneity due to age of evaluators (Margoni et al., 2021), evaluation methods (Sezer et al., 2016) and information controls (Brownback and Kuhn, 2019). Additionally, in recent years, there has been a growing number of empirical studies using observational data (Gauriot and Page, 2019; Kausel et al., 2019; Lefgren et al., 2015; Meier et al., 2022, 2023). These observational studies provide evidence of the outcome bias in a variety of environments, whereas the persistence and heterogeneity of the bias in fields are still unclear. In total, implications for solutions to the outcome bias are scarce, although it is a potential cost from the perspective of inefficiencies and inequities in the allocation of sanctions, rewards, and promotions in economic institutions (Gauriot and Page, 2019).

This study fills the above research vacuum by providing new evidence on the outcome bias. Furthermore, the findings make three key contributions to understanding the outcome bias. First, this study adds to the research on the outcome bias in the field, thus increasing the validity of this phenomenon. Previous studies examining the outcome bias in third-party evaluation settings utilize sports data. While these studies show the existence of the outcome bias in soccer and basketball, this study is the first research to show it in baseball. The evidence in another sport supports the argument that the outcome bias is a general tendency of decision-makers, at least in the context of sports, and is neither specific to a certain field nor a statistical artifact (Meier et al., 2023). In addition, the persistence of the outcome bias suggests that the problems of potential wage inefficiency and promotion inequality are even more serious.

Second, the findings of this study imply discrimination related to the outcome bias. The racial heterogeneity in the outcome bias may reflect the racial preferences of managers, given that approximately 90% of MLB managers are white. This discriminatory bias diverges performance indicators from true productivity in the sense that it induces racial majorities to play more games than similarly productive racial minorities. Then, a discriminatory wage structure arises in which racial minorities are paid less than whites, since annual salaries are determined based on their performance indicators. This type of discrimination has the econometric problem that estimates holding implicitly measured productivity constant understate its true magnitude (Parsons et al., 2011). Therefore, this study reminds caution in the interpretation of estimates identified as discrimination even using wage equations to fully control for performance indicators as a proxy of productivity, in such cases where there may be discrimination in the generating process of performance indicators.

Finally, this study provides insight into solving the problem of the outcome bias described above. The evidence from this study is consistent with previous research reporting that even highly specialized professional sports managers tend to be biased toward uninformative outcomes in decision-making (Gauriot and Page, 2019; Kausel et al., 2019). On the other hand, the new fact that bias decreases with experience adds two perspectives to the solution of the outcome bias. One is that the experience may make the managers aware of their biased decision-making. Another possibility is that experienced managers with sufficient information about players' skills have less outcome bias because they avoid heuristic decisions relying too heavily on outcomes. These complementary interpretations align with the explanation that the outcome bias has decreased due to the recent availability of more information and advanced analytics (Meier et al., 2023). They also suggest that decision-making can be improved through empirical learning and adequate information gathering.

The remainder of this paper is organized as follows. Section 2 reviews the principal-agent problem and provides the empirical strategy. Section 3 shows the benchmark result of the outcome bias and its persistence. Section 4 conducts further analysis to reveal the heterogeneity of outcome bias. Finally, Section 5 concludes the paper.

2 Identification strategy

2.1 Review of principal-agent problem

This section discusses outcome bias within the framework of the principal-agent problem based on Bolton and Dewatripont (2005). A principal evaluates an agent to set optimal compensation. Under hidden actions, the principal's evaluation of the agent depends on an outcome and a signal, both of which are affected by an action.

Consider an outcome is $q \in \{0, 1\}$, with $q = 1$ represents success and $q = 0$ represents failure. The agent chooses an action $a \in \mathbb{R}$ that positively affects the probability of success, defined as $\mathbb{P}(q = 1|a) = p(a)$ with $p'(a) > 0$. However, the action is not observable to the principal. Instead, the principal can observe a binary signal $s \in \{0, 1\}$, which is influenced by the agent's action. The joint probability of outcome and signal is given by $\mathbb{P}(q = i, s = j|a) = p_{ij}(a)$. Let w represent the compensation; the principal's utility function is then defined as follows:

$$V(q - w),$$

and the agent's utility function is defined as follows:

$$u(w) - a.$$

When both parties are risk averse, the principal offers compensation to the agent by solving the following:

$$\begin{aligned} &\underset{a, w_{ij}}{max} \sum_{i=0}^{1} \sum_{j=0}^{1} p_{ij}(a) V(i - w_{ij}) \\ &s.t. \underbrace{\sum_{i=0}^{1} \sum_{j=0}^{1} p_{ij}(a) u(w_{ij}) \geq 0}_{\text{Individual rationality constrains}}, \quad \underbrace{\sum_{i=0}^{1} \sum_{j=0}^{1} p'_{ij}(a) u(w_{ij}) = 1}_{\text{Incentive constrains}}. \end{aligned}$$

Let λ and μ be Lagrange multipliers of each constrains, then the optimal compensation w_{ij} respect following condition:

$$\frac{V'(i - w_{ij})}{u'(w_{ij})} = \lambda + \mu \frac{p'_{ij}(a)}{p_{ij}(a)}, \quad \lambda, \mu \in \mathbb{R}$$

The signal s drops out from the incentive scheme, when a change of action a yields the same rate of marginal change in the probability of each q. Similarly, the outcome is not incorporated into the optimal contract if there is a particular signal $s = \hat{j}$ that satisfies the following condition:

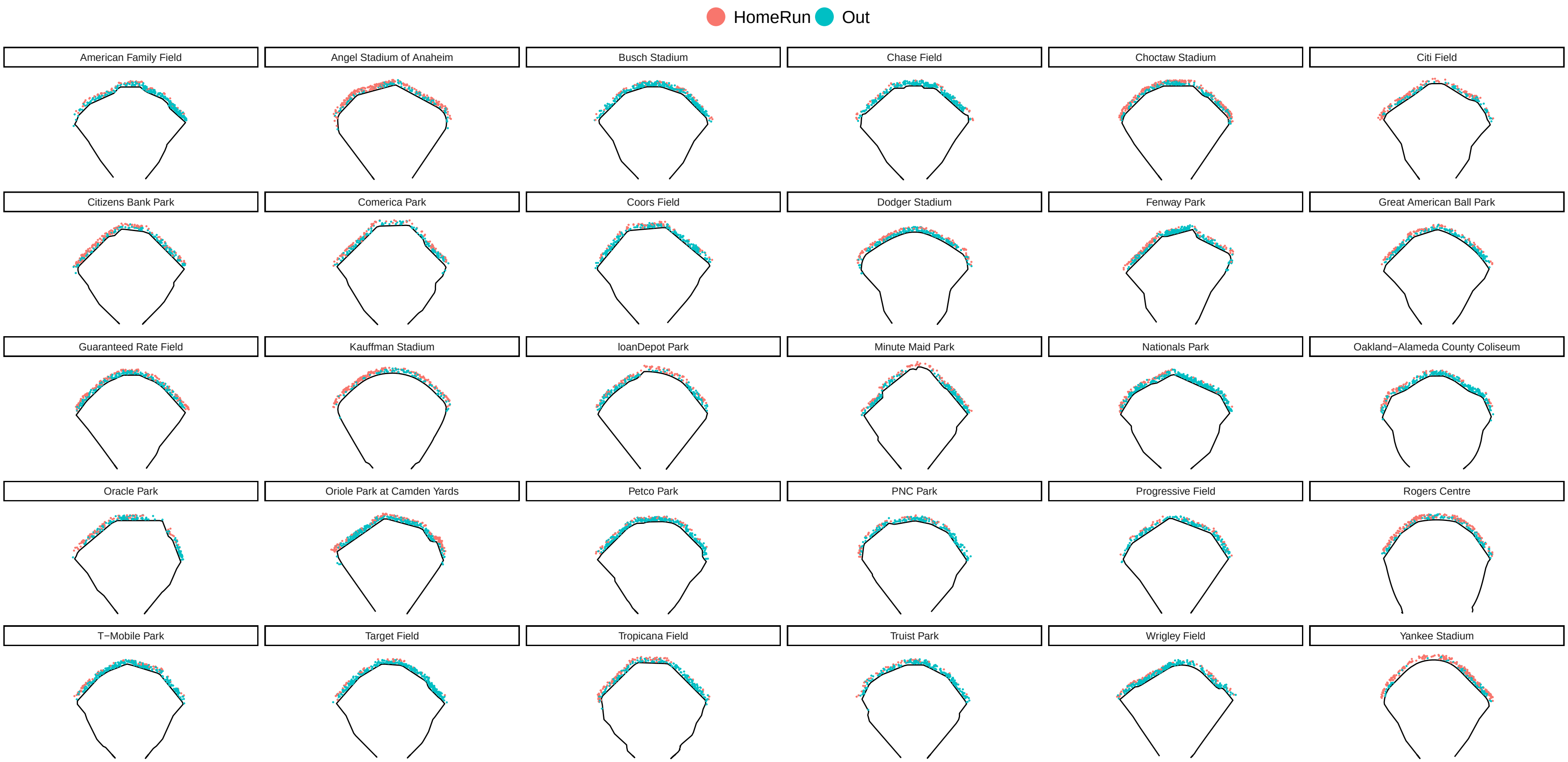
$$\frac{p'_{0\hat{j}}(a)}{p_{0\hat{j}}(a)} = \frac{p'_{1\hat{j}}(a)}{p_{1\hat{j}}(a)} \qquad \text{for } \ s = \hat{j}. \tag{1}$$

A sufficient condition to satisfy Eq (1) is

$$p'_{0\hat{j}}(a) = p'_{1\hat{j}}(a) = 0. \tag{2}$$

Eq (2) represents that the marginal effect of the action on success or failure is zero. In this case, outcome q is independent of the agent's action and should not work as information in the evaluation of the agent. Nevertheless, it is possible that if a randomly good (bad) outcome is obtained, the principal may consider the agent to have made more (less) effort and raise (lower) the evaluation. Such a deviation from the theory of optimal contracts can be considered as the outcome bias. The outcome bias can therefore be tested in situations where, conditional on an informative signal $s = \hat{j}$, the outcome q is as good as random in that it is not correlated with unobserved actions from the agent (Gauriot and Page, 2019). Based on this framework, this study extracts the situation in which the outcome is randomly determined, rather than dependent on the agent's actions, to identify the outcome bias.

Figure 1: Hits falling within 20ft near the wall



Source: The information of dimension of ballpark is derived from R package GeomMLBStadiums <https://github.com/bdilday/GeomMLBStadiums>. The information of landing point of a fly ball is derived from MLB Statcast database which is provided by Baseball Savant <https://baseballsavant.mlb.com/statcast_search>

Note: Most points indicating hits near the wall are plotted beyond the line because the line showing the outer frame of the ground is based on the root of the wall. Data from Braves home games in 2016 and 2015 were excluded because the Atlanta Braves moved to Truist Park in 2017. No data on hits from games at stadiums without ball-tracking radars was included in the sample.

2.2 Conceptual framework

To find situations where outcomes are effectively exogenous, this study leverages ball-tracking data from MLB. This data includes the trajectory of the hit as measured by radar installed at the stadium. From this data, I isolate cases where the outcome is not affected by the agent's control. Specifically, I extract hits near the wall, where marginal differences in distance determine whether a ball results in a homerun or an out—situations in which outcomes are plausibly random.

In baseball, where batters attempt to hit pitches exceeding 90 mph, they exert only limited control over the exact landing spot of the ball. A hit that clears the wall results in a homerun, while one that falls short may be caught for an out. Furthermore, even if the ball clears the wall, fielders can reach over and make a catch, further introducing exogenous variation. A difference of just a few feet can dramatically alter the outcome. Additionally, environmental factors such as wind and air pressure influence ball trajectories (Bahill et al., 2009), and ballpark-specific variations in fence height and outfield dimensions further contribute to outcome unpredictability. Given these conditions, it is reasonable to assume that hits near the wall represent a quasi-random assignment of outcomes, driven by external factors rather than player skill or strategy. Section 2.4 provides evidence to support this assumption.

2.3 Data

The main dataset is ball-tracking data derived from MLB Statcast database in 2015 - 2019. The data includes the coordinates of the location where each hit was landed, the name of the batter, the date and time, and the ballpark name. This dataset was merged with the dimensional data for each ballpark and limited to only the hits that fell within 20 feet of the fence. Figure 1 shows a visualization of the hits near the wall for each stadium. Of the 15,703 hits extracted, 8449 were homeruns and 7,254 were fly outs. The samples represent 18% of all homeruns and 3% of all fly outs, respectively. To clarify differences in outcome,

Table 1: Balance test of players' performance and characteristics

	Performance				Characteristics		
	HR	Out	Difference		HR	Out	Difference
Since the start of the season				Age	28.812	28.782	0.030
Plate appearance	203.045	206.515	−3.470	Experience	7.892	8.019	−0.127
Hit	53.837	54.884	−1.047	Catcher	0.101	0.094	0.007
Homerun	8.755	8.823	−0.068	Infielder	0.445	0.445	0.000
Batting average	0.256	0.257	−0.001	Outfielder	0.201	0.184	0.017*
Last season				White	0.520	0.528	−0.008
Plate appearance	391.706	397.691	−5.985	Black	0.216	0.212	0.004
Hit	103.675	105.629	−1.954	Hispanic	0.236	0.238	−0.002
Homerun	16.435	16.313	0.122	Asian	0.028	0.023	0.005
Batting average	0.257	0.258	−0.001				

Source: The information of performance statistics of last season and players' characteristics are derived from Lahman database <https://sabr.org/lahman-database/>. The information of performance statistics since the start of the season is calculated from the main dataset.

Note: The presented values are the mean of the covariates and the difference between them for players who hit homeruns near the wall and those who hit outs near the wall, respectively. Discrete variables for player characteristics are presented as proportions. The performance statistics since the start of the season show the cumulative performance for the season up to one game before the treatment. Batting average is determined by dividing a player's hits by their total at-bats. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

hits that did not clear the wall and were not directly caught by the fielder were excluded from the sample.

To test for the outcome bias, I also use the information on player line-ups for all games 2015 - 2019[1]. This data includes information on whether or not the player has played in a game and their batting order. The manager decides the line-up for each game, so it is possible to observe the manager's evaluation of the players in terms of giving them the opportunity to play. This data is connected to the main dataset to analyze the relationship between random outcomes and managers' decisions. In particular, the manager's decision in the next game, in which players have hit the ball to near the wall, is compared to a homerun and an out.

2.4 Balance test on selection bias

The key in identifying the outcome bias is whether the outcomes of hits near the wall can be assumed to be random. To the validity of this assumption, this subsection shows no

[1] The information is derived from Retrosheet <https://www.retrosheet.org/>.

difference in skills and characteristics between players who hit near the wall.

Table 1 shows the mean of the covariates and the difference between them for players who hit homeruns near the wall and those who hit outs near the wall, respectively. Of all the tests, none are significant at the 5% level. The left side of the table compares the players' performance indicators. For example, since the start of the season, the player who hit the homerun near the wall did not hit more homeruns than the player who was out (0.068, $p = 0.733$). The right side of the table compares the players' characteristics. Only the proportion of outfielders is significant at the 10% level, but all other covariates are not significant. There is no evidence of selection bias is found in these results, which supports the assumption that the outcomes are randomly determined.

3 Benchmark result

This section examines the managers' decisions against random outcomes. Three scales are used for the manager's decision: the probability of play the game from the start (Starter), the probability of play the game as batting order at 3 – 5 (Cleanup), and the probability of play the game as batting order at 1 – 4 (Top order). These probabilities are higher for better players and therefore work as a proxy variable for the managers' evaluation of the players.

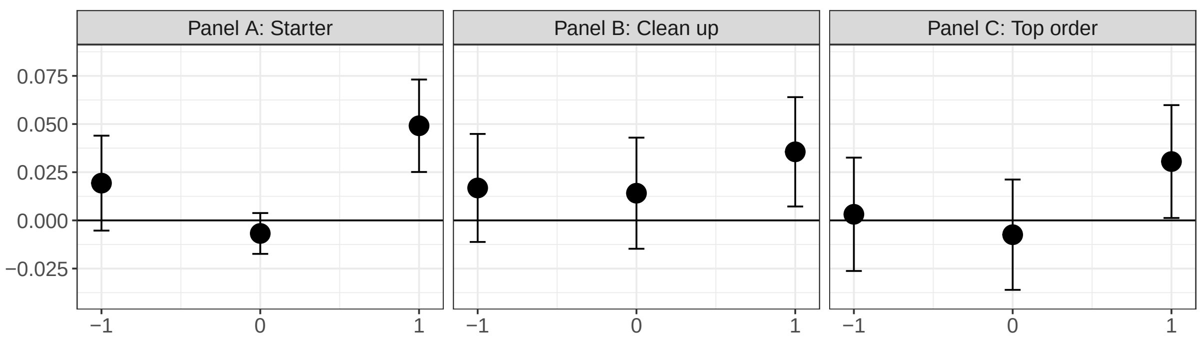
I conduct the regression analysis to test for the effect of hits near the wall on the managers' evaluation of the players. The linear probability model is extended to the local projection to check for the residual heterogeneity of players. Let i, t, and g denote the player, year, and game, respectively. The following equation is estimated:

$$w_{i,t,g+h} = \alpha + \beta_h Luck_{i,t,g} + \epsilon_{i,t,g} \tag{3}$$

where w is a dummy variable that takes the value of one for Starter, Cleanup, and Top Order, respectively, and zero otherwise; $Luck$ is a dummy variable that takes the value of one for a homerun and zero for an out. Eq (3) is estimated for $h \in \{-1, 0, 1\}$ at game frequency, where $h = 0$ means the game in which a player hits near the wall. Coefficient $\beta_{h=1}$ shows the outcome bias, $\beta_{h=0}$ and $\beta_{h=-1}$ for before the event provide the placebo test.

Figure 2 shows the estimation results. The horizontal axis denotes game frequency h and the vertical axis denotes coefficients β_h, the coefficients are nearly zero for term $h \leq 0$, which indicates no selection bias in the sampling procedure. In contrast, the estimates jump at $h = 1$, which is the game after hitting the ball near the wall. Using Panel A as an example, the player whose hit near the wall is a homerun is about five percentage points more likely to play as a starter in the next game than a player whose hit is an out. Similarly, the probability of playing as Cleanup and Top order is also high for players who hit homeruns near the wall. These results indicate that random outcomes influence managers' evaluation of players.

Figure 2: Identification of outcome bias and placebo test



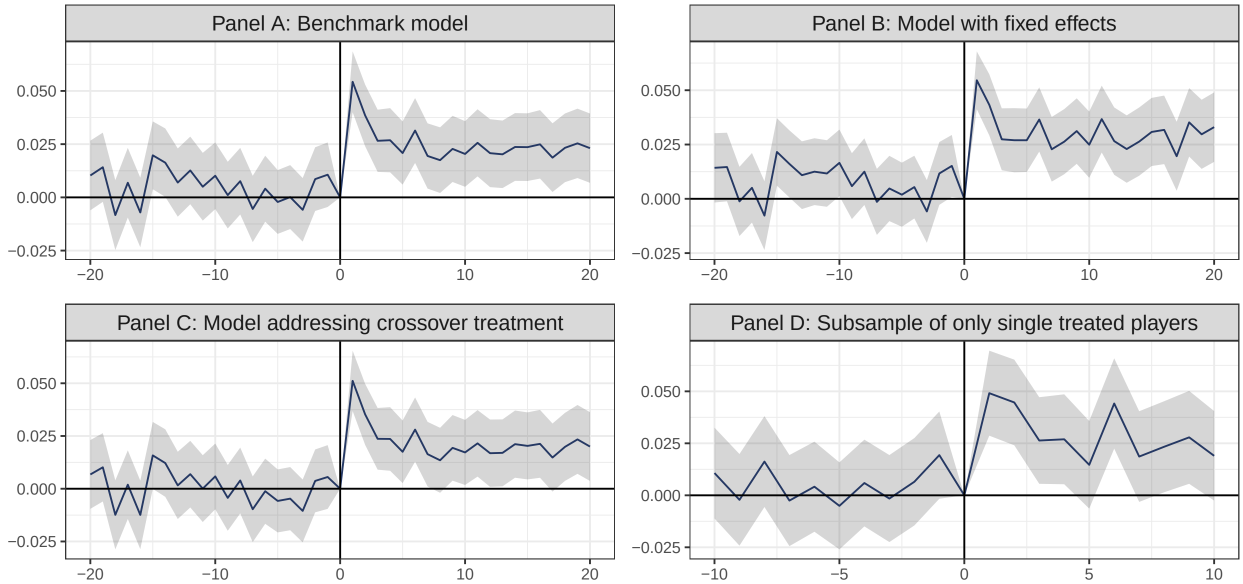
Note: Zero on horizontal axis indicates the game in which a player hits near the wall. The vertical axis represents β_h of Eq (3). The error bars denote 95% confidence interval.

4 Further analysis: persistence and heterogeneity

4.1 Persistence

This section examines how long the random outcome affects the managers' evaluation of the players. To test for the persistence of the outcome bias, the local projection is extended the term to 20 games before and after the event. I estimate β_h for $h \in \{-20, \cdots, 20\}$ of Eq (3). As the objective variable, the probability of Starter is used.

Figure 3: Persistence of outcome bias



Note: Each panel represents the impulse responses of the probability of playing as the starter to the homerun near the wall. Zero on horizontal axis indicates the game in which player hits near the wall. The vertical axis represents β_h of Eq (3). The shaded area denotes the 90% confidence interval.

Figure 3 shows the impulse responses of the probability of playing as Starter to the homerun near the wall. Panel A is the benchmark result. The estimate that jumped one game after treatment persists for at least 20 games while decreasing. Panel B shows the results with player fixed effects to control for unobserved heterogeneity. Impulse response is similar to panel A and the result is robust. These results indicate that the outcome bias decreases over time but persists.

Panels C and D are the result of addressing crossover treatment. Crossover treatment is an experimental design in which the same individual is assigned multiple treatments. In this experimental design, measurements may not only be affected by the treatment assigned most recently to a subject, but could also be affected by lingering effects of treatments that were assigned to the same subject in one of the previous periods (Stufken, 1996). In the case of this study, there may be a bias due to crossover treatment, as the players hit the ball near the wall several times during the observation period. To address this concern, I estimate the following equation:

$$w_{i,t,g+h} = \alpha + \beta_h Luck_{i,t,g} + \sum_{k=20}^{1} \gamma_k Luck_{i,t,g-k} + \epsilon_{i,t,g} \tag{4}$$

where $\sum_{k=20}^{1} \gamma_k Luck_{i,t,g-k}$ is lags of luck to capture the crossover effect of the previous $Luck$. Note that $Luck$ is normally a missing value for games with no hits near the wall, but in this term it is complemented by zero for estimation purposes. The estimation result of Eq (4) is shown in Panel C, and the estimates are consistent with the benchmark model. As a further robustness check for the crossover treatment, I estimate Eq (3) for $h \in \{-10, \cdots, 10\}$ using the subsample of only single-treated players during the analysis period. To ensure sample size, the analysis period is set to 10 games before and after the event. Panel D shows the result and estimates is consistent with the benchmark model.

4.2 Heterogeneity by players' race

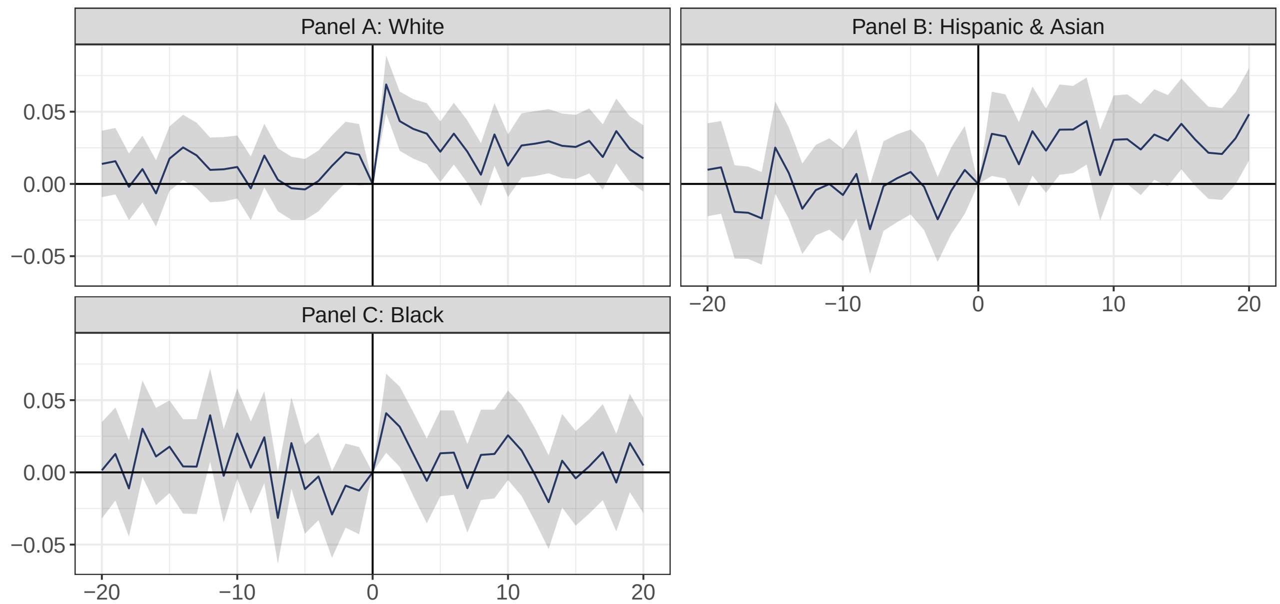
Table 2: Racial heterogeneity of outcome bias

	(1)	(2)	(3)
Luck	0.038^{***}	0.039^{***}	0.039^{***}
	(0.012)	(0.012)	(0.012)
White	-0.035^{***}	$N.A.$	$N.A.$
	(0.010)		
Luck × White	0.031^{*}	0.029^{*}	0.028^{*}
	(0.017)	(0.016)	(0.016)
Player FE		✓	✓
Team FE			✓
Observations	$8,864$	$7,093$	$7,093$

Note: Column (1) estimates Eq (5), and Column (2) estimates the model with the player fixed effect, and Column (3) estimates the model with the player fixed effect and the team fixed effect. Columns (2) and (3) do not estimate the coefficients for White due to player fixed effects. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

This section examines the racial heterogeneity of outcome bias. To begin, the players in the dataset are categorized as White, Black, Hispanic, or Asian. The classification method

Figure 4: Persistence of outcome bias by race



Note: Each panel represents the impulse responses of the probability of playing as the starter to the homerun near the wall. Zero on horizontal axis indicates the game in which player hits near the wall. The vertical axis represents β_h of Eq (3). The shaded area denotes the 90% confidence interval.

follows Parsons et al. (2011). Based on the birth country provided by the player's profile[2], I classify players born in Australia, Canada, and EU countries as White, players born in Hong Kong, Japan, South Korea, and Taiwan as Asian, and players born in Aruba, Bahamas, Brazil, Colombia, Cuba, Curacao, Dominican Republic, Honduras, Jamaica, Nicaragua, Puerto Rico, Panama, and Venezuela as Hispanic, respectively. Players born in other countries including US or without information on their birth country were visually inspected and classified based on their picture and family name. After this process, I reinspect the pictures of the players to finalize the classification sample.

First, I examine the racial heterogeneity of the effect of the hit near the wall on the probability of playing as a starter in the next game. Let $White$ be the dummy variable, which takes value of one for the white players and zero otherwise, and I estimate the following

[2] The information comes from the Larman database <http://www.seanlahman.com/>.

equation:

$$w_{i,t,g+1} = \alpha + \beta_1 Luck_{i,t,g} + \gamma White_i + \delta Luck_{i,t,g} \cdot White_i + \epsilon_{i,t,g} \tag{5}$$

where δ is the coefficient of the interaction term between $Luck$ and $White$ to capture racial differences in the outcome bias. Table 2 shows the estimation result. Column (1) reports the coefficients estimated by Eq (5). The coefficient of the interaction term between Luck and White is positively statistical significant. Column (2) reports the result of the model with player fixed effects to control for unobserved heterogeneity, and the results are robust. Column (3) reports the result of the model with team fixed effects to control for racial ideology specific to the city to which the team belongs, and the results are robust. These results indicates that the outcome bias in the managers' evaluations responds 2.9 $\sim$ 4.0 percentage points more strongly to white players than to players of other races.

Second, I compare the persistence of outcome bias by race. The dataset is divided into three groups: whites, blacks, and Hispanics & Asians, and the Eq (3) for $h \in \{-20, \cdots, 20\}$ is estimated. Figure 4 shows the estimation result. In Panel A, which reports the result of white players, the estimate jumps in the next game of the treated game and remains positive in subsequent games while decreasing. In Panel B, which reports the result of Hispanic and Asian players, the jump of the estimate are lower than for white players but remain positive. However, in Panel C, which reports the result of black players, the estimate that jumps after a treated game immediately drops to around zero. That is, the outcome bias against black players disappears more quickly than for other players. These results suggest that there is racial heterogeneity of outcome bias, not only in magnitude, but also in persistence.

4.3 Heterogeneity by managers' experience

This section examines the effect of the manager's experience on decision-making. To analyze the relationship between outcome bias and experience, the number of years working as a manager is used in estimation as experience. In addition, decisions to use players may

Table 3: Heterogeneity of outcome bias by manager's experience

Panel A: Model with the manager's managerial career

	(1)	(2)	(3)	(4)
Luck	0.0935^{***}	0.0885^{***}	0.0996^{***}	0.0837^{***}
	(0.0193)	(0.0176)	(0.0197)	(0.0179)
Experience	0.0010	0.0024^{*}	-0.0109^{*}	-0.0270^{***}
	(0.0011)	(0.0014)	(0.0057)	(0.0063)
Luck × Experience	-0.0030	-0.0025	-0.0048^{**}	-0.0030^{*}
	(0.0019)	(0.0017)	(0.0019)	(0.0017)
Player FE		✓		✓
Manager FE			✓	✓
Observations	$4,658$	$3,772$	$4,658$	$3,772$

Panel B: Model with the manager's age

	(1)	(2)	(3)	(4)
Luck	0.2302^{***}	0.2095^{***}	0.2579^{***}	0.2174^{***}
	(0.0846)	(0.0775)	(0.0865)	(0.0787)
Age	0.0008	0.0018	-0.0104^{*}	-0.0265^{***}
	(0.0009)	(0.0012)	(0.0056)	(0.0062)
Luck × Age	-0.0030^{*}	-0.0027^{*}	-0.0037^{**}	-0.0030^{**}
	(0.0016)	(0.0014)	(0.0016)	(0.0015)
Player FE		✓		✓
Manager FE			✓	✓
Observations	$4,658$	$3,772$	$4,658$	$3,772$

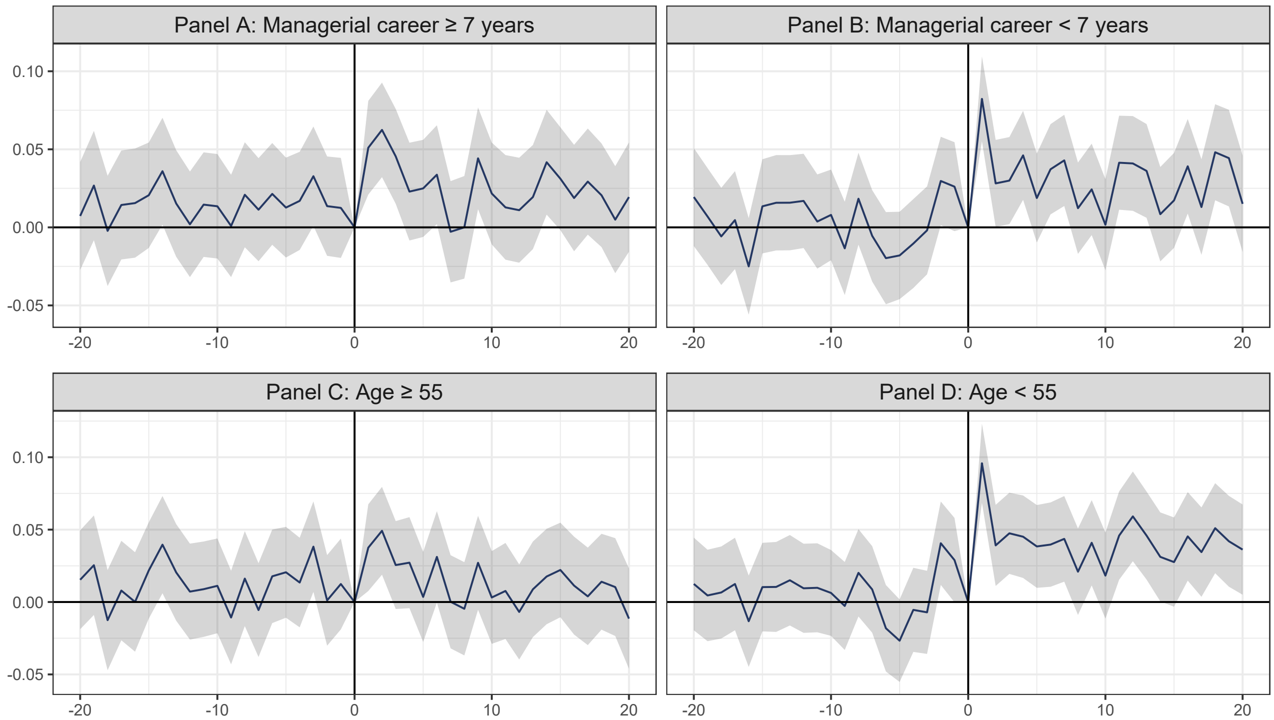
Note: Column (1) estimates Eq (7), and Column (2) estimates the model with the player fixed effect, Column (3) estimates the model with the manager fixed effect, and Column (4) estimates the model with the player fixed effect and the manager fixed effect. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

be influenced by their experience in coaches, staff, and other non-management roles. However, no information is available on the number of years worked as coaches or other staff. Then, the age of the manager is used as another proxy variable for experience. These new explanatory variables are each added to the estimation equation:

$$\begin{cases} w_{i,t,g+1} = \alpha + \beta_1 Luck_{i,t,g} + \gamma Exper_{i,t} + \delta Luck_{i,t,g} \cdot Exper_{i,t} + \epsilon_{i,t,g} & (6) \\ w_{i,t,g+1} = \alpha + \beta_1 Luck_{i,t,g} + \gamma Age_{i,t} + \delta Luck_{i,t,g} \cdot Age_{i,t} + \epsilon_{i,t,g} & (7) \end{cases}$$

where $Exper_{i,t}$ is the number of years of the managerial career of the manager in the team to which player i belongs in year t, and $Age_{i,t}$ is the age of the manager in the team to which

Figure 5: Persistence of outcome bias by age



Note: Each panel represents the impulse responses of the probability of playing as the starter to the homerun near the wall. Zero on horizontal axis indicates the game in which player hits near the wall. The vertical axis represents β_h of Eq (3). The shaded area denotes the 90% confidence interval.

player i belongs in year t. The interaction term between luck and each of the variables representing experience captures the heterogeneity of the outcome bias.

In Table 4.2, Panel A shows the estimation result of Eq (6). In Column (1), the coefficient of the interaction term between Luck and Experience is negatively statistical significant. Columns (2), (3), and (4) show models with player fixed effects, with coach fixed effects, and with both fixed effects, respectively. The estimates in these columns are almost consistent with Column (1). Panel B shows the estimation result of Eq (7). Also in Panel B, the coefficients of the interaction terms are negatively significant in all columns. These results suggest that the more experience the manager has, the less outcome bias there is in the decision. More specifically, the outcome bias decreases by 0.33 $\sim$ 0.49 percentage points for each additional year of the manager's managerial career and by 0.34 $\sim$ 0.40 percentage points for each additional year of the manager's age.

Additionally, I compare the persistence of outcome bias by managers' experience. The dataset is divided into two groups by experience and age, respectively, and the Eq (3) for $h \in \{-20, \cdots, 20\}$ is estimated. Panels A and B compare samples with a cutoff of 7 years of experience. At h = 1, the outcome bias in Panel A is greater than in Panel B, while the persistence is almost the same. However, there is a difference between Panel C and Panel D, which are the results for samples with a cutoff age of 55. In Panel C, which includes managers under age 55, the estimate jumps at $h = 1$ and remains positive thereafter. Conversely, in Panel D, which includes managers under age 55, the estimate is lower than that in Panel C and drops immediately to close to zero. This result suggests that experienced manager not only has a less outcome bias, but also less persistence. Thus, as managers gain experience, they may be able to reduce outcome-based decision-making. In addition, since the difference was more pronounced in the sample divided by age than by managerial career, it is possible that non-managerial experience, such as coaching or staffing, also contributes to improved their decision-making.

5 Conclusion and discussion

This study examines the outcome bias using high-frequency tracking data provided by MLB. The natural experimental setting extracted from the dataset allows this study to identify the outcome bias. The estimation results show that successful players in the random situation are highly evaluated by managers. The evidence of overreaction to outcomes supports the existence of the outcome bias as reported in other fields. In addition, the local projection analysis shows the persistence of the outcome bias. This new fact implies that the potential cost of outcome bias in economic institutions is more serious.

Further analysis sheds light on the racial heterogeneity of the outcome bias. The outcome bias is greater and its persistence is longer for white players. The heterogeneous outcome bias could be further magnified by the racial match (Parsons et al., 2011) between manager and player, although sample limitations with only 10% racial minority managers of the total precluded analysis in this study. This racially unbalanced outcome bias leads to a larger performance indicator for white players than the similarly productive other players. Then, the systematic distortion of performance indicators may cause a discriminatory wage structure. In this case, seemingly non-discriminatory salary assessment based on performance indicators may be disadvantageous to racial minority (Bohren et al., 2022). Furthermore, this type of wage discrimination should be attentive because it is under-estimated in standard econometrics (Parsons et al., 2011).

Finally, this study shows heterogeneous outcome bias by manager's experience. The outcome bias and its persistence decrease with increasing manager's experience. This result is partially inconsistent with the literature reporting that aging increases outcome bias (Margoni et al., 2021), whereas consistent with reporting that experience improves decision-making in finance (Breza et al., 2020; Campbell et al., 2014; Gine and Goldberg, 2023) and agriculture (Abdulai et al., 2023; Foster and Rosenzweig, 1995; Santeramo, 2018). The reduction of the bias may be because experienced managers notice it in their careers and correct their decisions. Alternatively, experience may allow them to avoid heuristic decisions that rely heavily on outcomes by gathering information about players' skills. In summary, the findings of this study suggest that decision-making can be improved through experiential learning and the collection of appropriate information.

References

Abdulai, A.-R., Quarshie, P., Duncan, E., and Fraser, E. (2023). "Is agricultural digitization a reality among smallholder farmers in africa? unpacking farmers' lived realities of engagement with digital tools and services in rural northern ghana." *Agriculture & Food Security*, *12*.

Alicke, M., and Davis, T. (2004). "On the robustness of outcome bias: No constraint by prior culpability." *Basic and Applied Social Psychology*, *26*, 131–146.

Alicke, M., Davis, T., and Pezzo, M. (1994). "A posteriori adjustment of a priori decision criteria." *Social Cognition*, *12*(4), 281308.

Bahill, A. T., Baldwin, D., and Ramberg, J. (2009). "Effects of altitude and atmospheric conditions on the flight of a baseball." *International Journal of Sports Science and Engineering*, *3*, 1750–9823.

Baron, J., and Hershey, J. C. (1988). "Outcome bias in decision evaluation." *Journal of personality and social psychology*, *54*(4).

Bohren, J. A., Hull, P., and Imas, A. (2022). "Systemic discrimination: Theory and measurement." *NBER Working Paper Series*, *29820*.

Bolton, P., and Dewatripont, M. (2005). *Contract Theory*, *MIT Press Books*, vol. 1. The MIT Press.

Breza, E., Kanz, M., and Klapper, L. F. (2020). "Learning to navigate a new financial technology: Evidence from payroll accounts." *NBER Working Paper Series*, (28249).

Brownback, A., and Kuhn, M. A. (2019). "Understanding outcome bias." *Games and Economic Behavior*, *117*, 342–360.

Campbell, J. Y., Ramadorai, T., and Ranish, B. (2014). “Getting better or feeling better? How equity investors respond to investment experience.” *NBER Working Paper Series*, (20000).

Foster, A., and Rosenzweig, M. (1995). “Learning by doing and learning from others: Human capital and technical change in agriculture.” *Journal of Political Economy*, *103*, 1176–1209.

Gauriot, R., and Page, L. (2019). “Fooled by performance randomness: Overrewarding luck.” *The Review of Economics and Statistics*, *101*(4), 658–666.

Gine, X., and Goldberg, J. (2023). “Experience in financial decision-making: Field evidence from malawi.” *Journal of Development Economics*, *161*, 103036.

Kausel, E. E., Ventura, S., and Rodrguez, A. (2019). “Outcome bias in subjective ratings of performance: Evidence from the (football) field.” *Journal of Economic Psychology*, *75*, 102132, applications of sports data to study decision making.

Lefgren, L., Platt, B., and Price, J. (2015). “Sticking with what (barely) worked: A test of outcome bias.” *Management Science*, *61*(5), 1121–1136.

Margoni, F., Geipel, J., Hadjichristidis, C., and Surian, L. (2021). “Inequity aversion in old age: An outcome bias in older adults’ socio-economic decisions.” *Cognitive Development*, *58*, 101016.

Meier, P. F., Flepp, R., and Franck, E. (2023). “Replication: Do coaches stick with what barely worked? evidence of outcome bias in sports.” *Journal of Economic Psychology*, *99*, 102664.

Meier, P. F., Flepp, R., Meier, P., and Franck, E. (2022). “Outcome bias in self-evaluations: Quasi-experimental field evidence from swiss driving license exams.” *Journal of Economic Behavior & Organization*, *201*, 292–309.

Parsons, C. A., Sulaeman, J., Yates, M. C., and Hamermesh, D. S. (2011). “Strike three: Discrimination, incentives, and evaluation.” *American Economic Review*, *101*(4), 1410–35.

Santeramo, F. G. (2018). “Imperfect information and participation in insurance markets: evidence from italy.” *Agricultural Finance Review*, *78*.

Sezer, O., Zhang, T., Gino, F., and Bazerman, M. H. (2016). “Overcoming the outcome bias: Making intentions matter.” *Organizational Behavior and Human Decision Processes*, *137*, 13–26.

Stufken, J. (1996). “Optimal crossover designs.” *Handbook of Statistics*, *13*, 63–90.