Toward an Understanding of Discrimination When Multiple Channels Exist*

Majid Ahmadi, Gwen-Jirō Clochard, Jeffrey Lachman, John A. List

Draft version: January 10, 2025

Abstract

When multiple forces potentially underlie discriminatory behavior, pinning down the precise sources becomes a challenge, making proposed policy solutions speculative. This study introduces an empirical approach, tightly linked to theory, to dissect two specific channels of discrimination: customer bias and managerial bias. To illustrate our framework, we integrate proprietary data with several publicly available datasets to uncover channels of discrimination within the Major League Baseball draft. Our analysis reveals that customer preferences significantly influence the drafting of players at the top end of the draft—those likely to gain immediate public attention and eventually play for the club. Conversely, we observe managerial homophily in the latter parts of the draft, where players who attract little attention and have minimal chances of playing for the club are selected. The observed preferential bias at both ends of the draft incurs a substantial opportunity cost. However, bias at the top end unduly affects competitiveness. Our findings provide significant implications for future research on measuring discrimination and addressing the challenge of multiple channels.

Keywords: racial bias, customer discrimination, sports

JEL Classification: J70, J71, M51

*Ahmadi: Georgia Institute of Technology, majid.ahmadi@gatech.edu; Clochard: Institute of Social and Economic Research, University of Osaka and Joint Initiative for Latin American Experimental Economics, gjclochard@osaka-u.ac.jp; Lachman: List Experimental Lab, University of Chicago, jeffrey.lachman@gmail.com; List: University of Chicago, Australian National University and NBER, jlist@uchicago.edu. We are grateful to Charlie Brobst, Charles Shi and Franco Albino for excellent research assistance. We thank Carlos Gomez-Gonzalez, Nobuyuki Hanaki and Marco Henriques Pereira for their helpful comments. We thank the participants of the Workshop on Economics of Institutions and Organizations at OSIPP and participants of the EEA-ESEM 2024 Meeting.

1 Introduction

Equitable treatment of citizens stands as a paramount issue for any nation, yet our grasp of the sources of marketplace discrimination remains speculative at best. This attribution problem becomes particularly thorny when multiple sources of discriminatory behavior intersect. Becker (1971) produced seminal work underscoring this complexity by identifying three potential biases that can drive discrimination (in hiring): managerial, co-worker, and customer biases. Over the past five decades, both theoretical and empirical literature have acknowledged these factors and worked to measure their import (Arrow, 1972; Phelps, 1972; Akerlof, 1997; Riach and Rich, 2002).

Despite strides in measuring and identifying discriminatory patterns (see, e.g., Jowell and Prescott-Clarke (1970); Goldin and Rouse (2000); Riach and Rich (2002); List (2004); Bohren et al. (2019, 2022)), a critical takeaway from the expansive literature on discrimination is the constraint imposed by data availability. This often compels researchers to rely on aggregated measures, thereby conflating the effects of diverse channels. Even field experiments might have difficulties controlling the assignment mechanism on several decision margins simultaneously. Yet, numerous facets of modern economies present scenarios where multiple discrimination sources coexist. Consider, for example, a restaurant employing both "front room" workers (e.g., bartenders, hosts, waitstaff) and "back room" workers (e.g., dishwashers, cooks, managers). Such settings offer a fertile ground to parse and understand the distinct nature of discriminatory preferences (Combes et al., 2016). This approach, theoretically, enables researchers to disentangle and examine potentially complementary and competing biases.

Our study commences with a foundational price theory model elucidating the various discriminatory channels prevalent in contemporary markets. Our theoretical framework posits that to effectively distinguish between sources of discrimination, three critical data sources, beyond recruitment data, are essential: 1) an objective metric for assessing worker quality, 2) detailed information regarding the composition of hiring decision-makers, and 3) a measure of customer bias magnitude. In our pursuit to identify a suitable setting for analyzing these channels, we identified the Major League Baseball draft as an ideal environment. Despite its specific nature, this market offers the precise variation needed to examine and quantify contributory channels of discrimination. This setup adheres to List's external validity Litmus Test for generating novel data: it is exceedingly challenging, if not impossible, to find a more appropriate setting, with available data, to test rigorously our multiple channel hypotheses (List, 2020, p.45).

To complete our analysis, we integrate proprietary data with several publicly available sources that we meticulously scraped from the internet, aiming to examine decisions made in the Major League Baseball (MLB) draft from 2008 to 2019. During this period, approximately 12,000 players were drafted by MLB teams. Our proprietary data, sourced from a partner MLB team, includes scouting evaluations and detailed information about each scout, including racial background. Publicly available data were gathered from multiple sources. First, we obtained data from the Perfect Game USA website,[1] which provides a list of approximately 60,000 players names and whether they were drafted in the MLB. Second, we accessed a comprehensive database of high school and college prospects drafted into the MLB, covering over 10,000 individuals with detailed characteristics and performance metrics. This dataset enables us to construct the first key metric necessary to distinguish sources of discrimination: an objective measure of player quality. Furthermore, we scraped data on all players who appeared in at least one MLB game from 2008 to 2019, sourced from www.espn.com. We subsequently categorized these players' racial affiliations using data from Rosenman et al. (2022).

We also collected comprehensive data on all scouting directors (those responsible for talent identification and recruitment strategy) across MLB teams, including their racial backgrounds. This represents the second critical dataset required to parse sources of discrimination: the racial composition of hiring managers. Using these data, we assess whether scouting directors exhibit a propensity to recruit players of their own race.

Our final data source is pivotal in enabling us to distinguish between sources of discriminatory behavior:

[1]Accessible here.

an external measure of customer (fan) bias. We leverage a naturally occurring event, the Black Lives Matter movement in June and July 2020, to construct this metric. Following the demonstrations, every MLB team posted messages on social media regarding Black Lives Matter, garnering significant attention (both negative and positive) from fans. This measure of fan bias comprises several components. We collected reactions to the teams' posts on Facebook and Twitter, utilizing the work of Kumar and Pranesh (2021) to analyze the message content. Employing their algorithm, which is based on approximately 9,000 manually labeled tweets related to Black Lives Matter, we created detailed information for each message. Additionally, as an alternative measure of customer bias, we analyzed stadium attendance data from 2008-2019, examining its correlation with the racial composition of the team. This multifaceted approach provides a comprehensive understanding of fan biases, crucial for parsing the different channels of discrimination.

Employing Mincer-type models to examine our dataset, we uncover several key insights. First, there appears to be no significant association between race and the likelihood of a player being drafted. In fact, when controlling for prospect quality, African-American players exhibit a slightly higher probability of being drafted compared to their White counterparts. Second, analyzing signing bonuses conditional on being drafted, we find that player compensation is generally consistent across racial groups. However, controlling for player quality, there is some evidence suggesting that Asian and White players receive lower signing bonuses compared to their peers.

Yet, this aggregation conceals a pivotal insight within our data, culminating in our third key finding: the data align with patterns of discrimination when we isolate individual channels while suppressing others. Specifically, empirical results reveal a strong correlation between the drafting of African-American players and fan bias during the early rounds of the draft: fan bias is associated with whether a player of a certain race is drafted early in the draft. This fan bias correlation, however, dissipates in the later rounds. Collectively, these findings suggest that MLB clubs are likely considering customer preferences when selecting players who will attract significant scrutiny and public attention.[2]

Our fourth major finding reveals homophily in the selections made by scouting directors across MLB clubs during the later rounds of the draft. Specifically, conditional on player quality, we observe that scouting directors demonstrate a bias toward players of their own race, with these players being approximately 38 percent more likely to be drafted in rounds 26-40. This finding aligns with the notion that, when the stakes are lower and public scrutiny is reduced, scouting directors are more inclined to express their personal preferences. This is supported by the fact that players drafted in these rounds have a very low probability of reaching the MLB, with less than 10 percent making it in our sample period.

Our final result indicates that the expression of such biased preferences, whether driven by customers or managers, carries certain economic costs. For instance, our estimates suggest that teams are willing to draft players valued at approximately $1 million less when fan bias increases by one standard deviation. However, when examining scouting biases, the financial impact is less significant because the players they draft, although less likely to make the MLB club than their peers, are still long-shots to reach the majors.

Even though our findings emerge from a specific segment of the world, the broader lessons drawn from our study are of significant importance. First, a key methodological insight is that researchers analyzing discrimination with establishment-level data must be cautious of the risks associated with aggregate measurements. If there is a tension between biased preferences of management and customers, aggregate data can underestimate or obscure key biases. This risk is heightened when trade-offs exist across different decision-making margins. Second, our results underscore the importance of relative prices in the context of discrimination. An organization may prioritize catering to customer preferences, but as the rewards for such bias diminish, internal organizational preferences gain prominence. Conversely, while homophily plays a role, it is not strong enough to manifest when the stakes are high.

Finally, from a policy perspective, it is crucial to pinpoint the exact channels of bias to develop effective

[2]It is important to note that we do not imply causality here. The causality could indeed flow in the opposite direction, where fan biases are influenced by the under-representation of certain races on their preferred teams.

and scalable interventions. Our study offers a framework for modeling and estimating the relevant sources of discrimination, providing insights that can be generalized to other settings. Additionally, our methodological approach illustrates the utility of standard economic models in identifying which markets and environments are most suitable for exploring these issues.

The remainder of our study is organized as follows. Section 2 includes a brief survey of the literature on customer-side versus supplier-side discrimination. Section 3 outlines a price theory framework to highlight the margins of choice. Section 4 includes the data used and Section 5 summarizes the empirical strategy. 6 discusses the key empirical results. Section 7 concludes.

2 Literature Review

In his seminal work, Becker (1971) identifies three sources of bias that contribute to discrimination: managers (in hiring decisions), co-workers, and customers. This paper specifically addresses discrimination stemming from managers (represented by the scouting team) and customers (the fans). Additionally, it contributes to the body of literature that applies these principles within the realm of sports. In this section, we provide a brief summary of these literature strands and elucidate how our current research intersects with and builds upon them.

2.1 Biased managers and co-workers

The literature on biased managers has predominantly focused on documenting discrimination in hiring, often utilizing correspondence studies as a methodological approach (Riach and Rich (2002); List (2025)). The evidence indicates that racial discrimination remains a pervasive issue in the U.S. labor market and is not showing signs of abating over time (Quillian et al., 2017). Instances of discrimination have also been identified in various other domains, including housing (Auspurg et al., 2019), transportation (Liebe and Beyer, 2021), online dating (Potârcă and Mills, 2015), clothing retail (Bourabain and Verhaeghe, 2019), childcare (Boyd-Swan and Herbst, 2019), law enforcement interactions, such as speeding tickets (Aggarwal et al., 2022) and even charitable giving (Landry et al., 2006).

This body of literature is intimately connected to the study of homophily within the labor market. Research in this area has consistently demonstrated that individuals tend to form groups that are homogeneous in terms of age, gender, religion, ethnicity, and even behavior and ability (Bell, 1981; McPherson et al., 2001). In the labor market, Giuliano et al. (2009) found that, within a specific supermarket chain, White managers tend to recruit fewer African-American workers compared to African-American managers, although they could not determine whether this disparity arises from supply or demand factors. Similarly, Benson et al. (2024) observed that managers in a U.S. retail context tend to hire workers from their own racial group. Our study underscores that homophily can endure even within highly competitive environments, particularly when public scrutiny is limited.

Our study also contributes to the literature on biased managers and performance. Glover et al. (2017) demonstrate that the productivity of minority workers diminishes under biased managers. Bandiera et al. (2009) further reveal that workers are more productive when they have connections to managers, particularly when the managers' compensation is fixed. In the medical literature, FitzGerald and Hurst (2017) provide a comprehensive review showing that doctors' biases towards patients from their own groups are correlated with implicit biases and quality of care. In our study, we find that players drafted by scouting directors of their own racial group do not exhibit lower productivity, as measured by performance, except for those who are low productivity employees. Furthermore, Lippens et al. (2021) show that penalties for discriminatory behavior are inversely related to the occurrence of such behavior. Similarly, our findings indicate that scrutiny of actions, as measured by draft round, is negatively correlated with biased decisions by scouting teams. The race of the scouting director does not influence draft decisions in the early rounds.

The literature on racial bias has also extensively explored the relationship between biased coworkers and both productivity and retention. For instance, Giuliano et al. (2009) find that employees are less likely to quit when a greater proportion of their coworkers share their racial background, particularly among White and Asian workers. In addition, Bygren (2010) similarly observes that workers are *less* inclined to leave when the share of the opposite gender is larger. Hedegaard and Tyran (2018)demonstrate that workers are willing to sacrifice 8% of their earnings to work alongside someone of the same ethnicity. Moreover, Hjort (2014) and Hamilton et al. (2012) reveal that team diversity can reduce productivity in settings such as Kenya and California. In our paper, however, due to the absence of data on coworker (in this context, other players) bias, we are unable to investigate the link between coworker bias and labor market outcomes. Consequently, we also cannot examine whether the presence of biased players within a team impacts draft strategies.

2.2 Biased customers

The present paper also contributes to the literature on the impacts of customer bias. Much like the bias exhibited by managers, there is substantial evidence indicating that some customers hold biases against certain groups of workers (Lang and Spitzer, 2020). For example, (Doleac and Stein, 2013) find that Black online sellers receive fewer and lower offers compared to White sellers. Similarly, Zussman (2013) identifies discrimination against Arab sellers of used cars in Israel, while Parrett (2011) and Kelley et al. (2024) present evidence of bias against waitresses and female online sales agents, respectively. Although previous studies have underscored the existence of customer bias, our paper takes this a step further by employing an external measure of customer bias derived from reactions to Black Lives Matter posts. This allows us to examine the degree of substitutability between a worker's race and their quality.

There is also compelling evidence linking customer bias to firms' hiring strategies. Bar and Zussman (2017), for instance, shows that a significant portion of customers in Israel discriminate against Arab workers, leading firms to reduce their employment of Arab workers and offer significantly lower service prices for those who do. Boyd-Swan and Herbst (2019) similarly find a strong correlation between the proportion of minority children in a childcare center and the likelihood of a minority applicant being selected.

Interestingly, Kuppuswamy and Younkin (2020) that film industry employers discriminate against minority actors due to perceived public bias, despite evidence that films with diverse casts actually generate higher revenue than those with homogeneous casts. In our study, we show that teams with fans who exhibit greater bias against African-Americans tend to recruit fewer such players. Additionally, customer bias has a more pronounced influence on firms' hiring strategies for highly productive employees.

In related work, Holzer and Ihlanfeldt (1998) demonstrate that the racial composition of an establishment's customer base closely correlates with the race of hired workers, particularly in roles requiring close customer interaction. Combes et al. (2016) also find that an increase in French residents exacerbates the wage gap between French and African workers with customer contact, more so than in sectors without such interaction. Laouenan (2017) further provide evidence of discrimination from both employers and customers against African-American entry-level workers in the US.

Regarding the performance effects on firms catering to biased customers, Leonard et al. (2010) show that in areas with larger White populations, higher numbers of African-American employees correlate with smaller sales. They also illustrate that aligning employee demographics with those of potential customers yields minimal benefits, except in cases where customers do not speak English. Conversely, Avery et al. (2012) identify a positive correlation between the racial and ethnic representativeness of the workforce and productivity, especially in stores with larger minority customer bases.

We contribute to this line of work by showing that there is little evidence that teams' draft strategies are driven by their local demographic composition. However, we observe that teams with biased fanbases are more inclined to recruit lower-quality White players to align with their fans' biases.

2.3 Racial bias in sports

In this section, we review the literature on racial bias within sports. Owing to the abundance of accessible data, sports have long been a fertile ground for testing economic theories (Kahn, 2000). Significant evidence highlights biases against racial minorities and immigrants in sports, ranging from NFL broadcast commentary (Schmidt and Coe, 2014) and European soccer (McLoughlin, 2020; MacLeod and Newall, 2022; Principe and van Ours, 2022) to amateur clubs recruiting fewer immigrants (Gomez-Gonzalez et al., 2021), and bias in the eligibility of international players in MLB (Hauptman, 2009; Ross and James Jr, 2015). Our study advances this literature by disentangling two potential sources of bias in drafting minority players: management bias (scouts and scouting directors) and fan bias.

Customer bias in sports is also well-documented. For instance, Kahn and Sherer (1988) provide evidence of attendance biases against African-American players in the NBA, with White players significantly boosting attendance by 5,700-13,000 per year per White player during the 1980s, although subsequent studies due to Dey (1997); McCormick and Tollison (2001); Kahn and Shah (2005) did not confirm these findings. Similarly, Brown and Jewell (1995) show that White college basketball players generate substantially more revenue than African-American players, creating a strong incentive for colleges to favor White recruits. Kanazawa and Funk (2001) further demonstrate that White players significantly increase Nielsen television ratings.

To identify evidence of customer discrimination, some studies focus on contexts where teams make no decisions, such as memorabilia markets or online player evaluations. For instance, Broyles and Keen (2010); Bryson and Chevalier (2015) find no aggregate racial bias in NBA trading cards and the Fantasy Premier League, respectively. Sur and Sasaki (2020) also find no evidence of bias in Indian cricket player evaluations. However, racial discrimination in the baseball card market is evident, with Nardinelli and Simon (1990) showing that cards for non-White players are priced lower than those for White players, and List (2004) finding that minority sellers receive significantly lower offers. On the contrary, Depken II and Ford (2006) report that African-American and Hispanic players receive more favorable votes for the All-Star Game ballot compared to White players. In this paper, we leverage an external event—-teams posting Black Lives Matter messages—to measure customer bias, helping us untangle the relative impact of fan bias on firms' hiring decisions.

Similar to the general literature on customer bias, several studies in this area document a positive correlation between the racial composition of basketball teams and their metropolitan areas (Brown et al., 1991; Burdekin and Idson, 1991; Hoang and Rascher, 1999; Burdekin et al., 2005). However, Kahn and Sherer (1988); Kahn and Shah (2005) find no evidence of racial bias in the NBA draft, controlling for college performance. Our findings contrast with these results, as we find no link between metropolitan composition and team draft strategy but evidence of racial bias in recruiting, controlling for prospect quality.

A few studies have also examined the evolution of fan bias in sports over time. Kahn (2009) reports a decline in racial discrimination from NBA fans since the 1980s, while Maennig and Mueller (2021) show an increase in fan bias in attendance in the 2000s compared to the 1980s. Our paper documents a stronger link between scouting directors' race and draft decisions in the early years of our data (2008-2015).

Having data on both management bias (scouting directors recruiting players from their racial group) and customer bias allows us to pinpoint customer preferences more precisely. This enables us to demonstrate that players' race and quality are complements in the fans' utility function.

3 Theoretical Framework

In this section, we introduce a simple price theory model to elucidate how various sources of bias can influence drafting decisions for different players. The model also serves to highlight the necessary data to populate empirically the various sources of bias. In this manner, it provides the theoretical underpinnings for why the context of the baseball draft, despite its seemingly special nature, serves as an exemplary setting for disentangling multiple sources of discriminatory behavior and illuminating the fundamental aspects of the issue.

3.1 Model setup

Setup A firm chooses to recruit one of two workers, W and B, with quality q_W and q_B. For simplicity, we assume the workers' productivity to be perfectly observable by all actors (managers and customers).

Customers Customers gain utility from the quality of the worker but also from the worker group membership. We assume consumers have a constant elasticity of substitution (CES) utility function between race and quality of the worker. If the firm recruits worker W, the consumers gain the following utility:

$$U_W = [q_W^\rho + \beta^\rho]^{\frac{1}{\rho}} \tag{1}$$

β represents the degree of bias from consumers, and ρ represents the degree of substitutability between group membership and worker quality ($\rho < 0$ corresponds to complements, $\rho > 0$ corresponds to substitutes). We assume that firms know their customers' type.

In the empirical analysis, we assume that the parameter β corresponds to bias against African-American workers. If the firm recruits worker B, the consumer utility is

$$U_B = q_B \tag{2}$$

Demand and profit functions We assume a demand function that is linear in consumers' utility:

$$D = A - B(P - U) \tag{3}$$

The firm's profit from hiring worker $i \in W, B$ is then given by the following, where w_i is the wage of worker i:

$$\Pi_i = P \times D_i - w_i = P(A - B(P - U_i)) - w_i \tag{4}$$

Optimal strategy The firm chooses the worker type that maximizes profit, and therefore recruits worker W if $\Pi_W > \Pi_B$, which translates to

$$BP(U_W - U_B) \geq w_W - w_B \tag{5}$$

Managers' decisions We also assume that managers make the hiring decisions on behalf of the team, in a principal-agent framework. They desire to maximize firm's profit, but also have utility over the identity of the worker. As such, their utility functional for recruiting a worker is given by:

$$u_i^M = \Pi_i + \frac{\gamma}{q_i} \times \mathbb{1}(W) \tag{6}$$

where γ is the bias parameter of recruiting managers. Assuming the functional form $1/q$ means that quality of workers reduces the bias in favor of W. The functional form represents higher scrutiny from owners for the recruitment of highly-productive employees, representing a potential wage penalty for imposing their own views (Lippens et al., 2021). An alternative modeling approach that produces a similar functional form would be to assume that managers' decisions impact the firm's profit only for workers visible to customers. This can be illustrated by considering restaurants that hire both waitstaff (customer-facing) and dishwashers (non-customer-facing). Previous research, for example, has demonstrated that discriminatory behavior is more prevalent in roles involving direct customer interaction. (Combes et al., 2016). In the empirical analysis that follows, we posit that managers exhibit a preference for workers of their own racial group and systematically test this hypothesis.

3.2 Predictions

Effect of manager bias On the management (internal) side, the functional form of the utility function of managers is decreasing in the workers' quality, suggesting that their bias will be stronger for low productivity workers. This leads to our first prediction.

Prediction 1: Managers are more likely to recruit players from their own race for low quality workers.

Effect of customer bias If consumers have a strong preference for worker W (i.e, a higher β), then the difference in utility will be higher (Equation 1), and the firm will be more likely to hire the W worker. This leads to second prediction.

Prediction 2: Firms with more biased customers will recruit fewer African-American workers.

Complementarity and quality When worker quality and race are complements ($\rho < 0$), the difference in utility between W and B becomes more pronounced with higher quality for a given β. This implies that the bias in favor of worker W intensifies for high-quality workers compared to low-quality workers. Conversely, if quality and race are substitutes ($\rho > 0$), then the utility difference decreases with quality. This means the bias in favor of worker W diminishes for high-quality workers.

Therefore, the variation in the effect of bias on quality is highly contingent upon the degree of substitutability between worker race and quality (Equation 5). It is therefore crucial to test whether fan bias exerts a stronger influence on high- or low-quality workers to discern if race and quality are substitutes or complements. As a result, we have a third prediction.

Prediction 3 (case of complements): A stronger correlation between customer bias and the firm's hiring decision for high-quality workers would suggest that customers perceive worker race and quality as complements.

3.3 Using the baseball draft to pin down fundamental parameters of the model

Our theoretical framework offers a comprehensive yet simple model for understanding decision-making within firms that encounter biases from both customers and managers. This framework is versatile and applicable across various industries, where different types of workers interact with customers and managers in distinct ways, leading to diverse biases.

For instance, in a restaurant setting, customer discrimination against certain groups of workers might result in the exclusion of those workers from waiter positions. Conversely, customer bias would not necessarily deter managers from hiring these individuals for backroom roles, such as dishwashing or cooking. Similarly, in supermarkets, customers interact primarily with cashiers rather than shelf stockers, while managers likely have more interactions with stockers than with cashiers. If managers are biased, they might avoid hiring certain workers for these roles. However, attributing the composition of the workforce in different occupations solely to customer or managerial bias is challenging in the general labor market, as selection into various roles could also influence workforce composition.

To evaluate the relevance of our model and comprehend the structure of customer preferences, our model provides several pieces of information that are essential in the empirical estimation: 1) an objective measure of worker productivity, 2) data on the racial composition of hiring managers to identify potential biases towards their racial group, and 3) an external measure of customer bias (β parameter). In the general labor market, obtaining such comprehensive data is quite challenging. However, the baseball draft presents a unique opportunity by offering a pool of comparable players all vying for the same occupation, thereby facilitating a more precise analysis of these parameters.

Furthermore, professional sports provide an abundance of publicly accessible data, which is invaluable for testing economic and managerial theories (Day et al., 2012; Fonti et al., 2023; Glennon et al., 2024). Specifically, baseball offers extensive performance data *before* players are drafted, including their high school and college performance, as well as information on players once they reach the major leagues. The draft round system serves as a reliable indicator of player quality, given that the probability of playing in the majors is strongly correlated with draft round. Additionally, professional sports facilitate the collection of data on high-level management, particularly scouting directors, who are responsible for identifying top prospects and developing recruitment strategies.

The baseball context also provides a unique advantage by allowing us to move beyond merely identifying customer bias and delve into customer preference functions. Within an industry, it is crucial to measure the degree of customer bias, and to understand how it is linked to discriminatory behavior. While obtaining such data in general economic settings is challenging, this paper leverages an external shock—the Black Lives Matter movement—to measure bias in a way that is not directly linked to performance. This approach enables us to determine whether preferences for player quality and race function as substitutes or complements.

4 Data

Our simple model highlights the rich data demands that are necessary to explore the various channels of potential bias. To add empirical content to the model, we use several sources of data. Each is necessary to analyze sources of racial bias in the baseball draft. We begin with the players themselves, and provide descriptive statistics in Table A.1.

4.1 Data on all players

The first dataset is sourced from the Perfect Game USA website,[3] which lists approximately 60,000 player names and indicates whether they were drafted into Major League Baseball. We combine the players' last names with data from Rosenman et al. (2022) to estimate the most likely race of each player. This racial measure is imprecise, so we refine it for a subset of drafted players, using self-identification data. The website also provides player evaluations (PG Grade). Since this prospect quality metric is imperfect—scout evaluations may be influenced by player race—we use our own measure of player quality when possible (see below). These data are necessary for the extensive margin analysis below.

[3] Accessible here.

4.2 Data on MLB drafted players

The second data source is a comprehensive database of high school and college prospects who were drafted into the MLB. This dataset includes over 10,000 prospects, detailing their characteristics and performance, provided by a partner MLB team. This team has collected racial self-identification data for approximately 2,000 players. Additionally, we enhance this dataset with a measure of prospect quality derived from a machine learning algorithm developed in a prior study (Ahmadi et al., 2022). This algorithm evaluates players' high school or college performance and is trained on the outcomes of players who successfully made it to the MLB. This machine learning-based metric serves as our measure of prospective player quality within the theoretical framework.

In addition to data on drafted players, we scraped information on all players who played at least one game in the MLB from 2008-2019.[4] We then categorized their racial affiliation using the racial identification data from Rosenman et al. (2022). This dataset includes players who were never part of the draft but were recruited as free agents, which increases the proportion of Hispanic and Asian players compared to the draft data.

4.3 Data on scouting teams

The third data source encompasses the entirety of scouting directors for all MLB teams, including their racial backgrounds. This database is pivotal in assessing whether the race of scouting directors influences draft strategies and if they exhibit a tendency to recruit players from their own racial group. The fourth data source, provided by a partner MLB team, includes detailed scouting evaluations and information about each scout, including their race. This dataset enables us to examine whether scouts display a preference for evaluating players of their own race more favorably, while controlling for player quality.

[4] Gathered from the ESPN website.

4.4 Data on fan bias from Black Lives Matter

Our fifth data source is a metric of fan bias, derived from the Black Lives Matter movement in June and July 2020. Following the demonstrations, every MLB team posted messages about Black Lives Matter on social media, which garnered various reactions from fans. Our measure of fan bias integrates several elements: i) Social Media Reactions: we collected reactions from team posts on Facebook and Twitter, ii) Content Analysis: utilizing data from Kumar and Pranesh (2021), we analyzed the content of these messages. They employed an algorithm trained on approximately 9,000 tweets related to Black Lives Matter, manually labeling each tweet to create a dummy variable (1 if the message is flagged as negative), iii) Comment Bias: we defined bias in comments as the proportion of negative comments on the team's post, and iv) Reaction Ratios: we also gathered the ratio of "angry" reactions to "likes" on Facebook posts.

For replies to tweets, restricting responses to followers of the team's Twitter page provided similar metrics (detailed in Table B.1). To synthesize these metrics, we calculated the sum of the three normalized elements, creating a composite Index of fan bias. The specific measures for each team are presented in Table B.3.

4.5 Alternative measure of customer bias: stadium attendance

The sixth data source pertains to stadium attendance for every game played between 2008 and 2019. We use these data to derive an alternative measure of fan bias. The process involves several steps. We first normalize attendance figures at the team level. Next, we calculate the average attendance on a monthly basis. We then regress the normalized attendance on the share of African-American players who participated in at least one game each month, controlling for the share of wins in that month and including year-fixed effects. For simplicity and to ensure comparability with the index variable defined earlier, we normalize the value of this variable and take the negative of the coefficient. Thus, the variable "Bias attendance" is positive if, for a given team, increasing the share of African-American players corresponds to fewer people attending the games. The attendance bias data is presented in Table B.4.

4.6 Other data sources

For the external side analysis, we complement the analysis by including the following controls: racial and political composition of fans ("Fans Black" and "Fans Republicans"), using the data from a Morning Consult Brand Intelligence survey on racial and political composition of each major U.S. sports teams (Silverman, 2020); racial and political composition at the local level, with the share of African-Americans ("Local Blacks") and the county-level vote share for Donald Trump in the 2020 presidential election ("Local Republicans").[5]

5 Empirical Strategy

5.1 Extensive margin analysis

We begin by analyzing the probability of players being drafted into the MLB, examining the influence of their race and quality (as determined by their highest ranking on the Perfect Game website). The race of the players is determined using the most likely race inferred from their last names, cross-referenced with data from Rosenman et al. (2022)). We estimate the following linear probability model:

$$\mathbb{P}(\text{Drafted}_i = 1) = \beta_0 + \beta_R \text{ Race}_i + \gamma \text{ PG Grade}_i + \epsilon_i \tag{7}$$

where Drafted is a binary variable indicating whether the player was drafted into the MLB, Race is a categorical variable representing the player's race, and PG Grade is the evaluation given to players by the Perfect Game website, serving as a proxy for player quality.

After examining the relationship between race and the probability of being drafted, we turn our focus to analyzing the signing bonus amounts, conditional on player quality. For this analysis, we consider players

[5] For the New York Mets and New York Yankees, we took the local shares of Republicans and African-Americans for all of New York City. For the Washington Nationals, we took the values for Washington, D.C. There is no value for the vote for Donald Trump for the Toronto Blue Jays, the team being located in Canada.

for whom signing bonus information is available (approximately 1,100 players). We estimate the following linear equation:

$$\text{Bonus}_i = \beta_0 + \beta_R \text{ Race}_i + \gamma \text{ ML quality prediction}_i + \epsilon_i \tag{8}$$

5.2 Management bias analysis

To assess bias within teams' management, we conduct two analyses. First, we evaluate whether players are more likely to be drafted by a team if they share the same racial group as that team's scouting director. For this, we use our dataset of drafted players to create a dataset with 30 observations per player (one for each team). We include two dummy variables: one indicating whether the player shares the same race as the scouting director, and another indicating whether the player was drafted by the team. We then estimate the following linear probability model:

$$\mathbb{P}(\text{Drafted}_{i,j} = 1) = \beta_0 + \beta_R \text{ Race}_i + \gamma \text{ X}_i + \delta \text{ Same race}_{i,j} + \epsilon_{i,j} \tag{9}$$

In Equation 9, we control for prospect quality as well as the distance to the team and the scouting director, a factor previously shown to influence the probability of being drafted by a specific team (Ahmadi et al., 2022). The coefficient of interest is δ. We estimate the model for the entire sample, and subsequently divide the sample based on draft rounds: rounds 1-25, and rounds 26+, anticipating from our model that the effects of internal bias will be more pronounced in the later rounds (Prediction 1).

The second dimension of our internal bias analysis focuses on the evaluation of players by scouts. Using data from a partner MLB team, we aim to analyze the link between a player's race and the scout's race. We estimate the following linear equation, with η being the coefficient of interest:

$$Evaluation_{i,j} = \beta_0 + \beta_R \text{ Race}_i + \gamma_R \text{ Race}_j + \delta \text{ ML quality prediction}_i + \eta \text{ Same race}_{i,j} + \epsilon_{i,j} \tag{10}$$

The bias in favor of scouts' own race could potentially be used as a measure of management bias (γ

parameter in Equation 6), but unfortunately we cannot have access to this measure for all teams.

5.3 Customer bias analysis

For the external bias analysis, we replicate the methodology used for the internal bias by creating a dataset with 30 observations per player and integrating it with the level of bias from the teams. Given that our measure of fan bias is defined as a bias against African-Americans, we focus primarily on the relationship between fan bias and the drafting of African-American players. We then estimate the following equation, with η being the main coefficient of interest.

$$\mathbb{P}(\text{Drafted}_{i,j} = 1) = \beta_0 + \beta_R \text{ Race}_i + \gamma \text{ X}_i + \delta \text{ Fan bias}_j + \eta \text{ Black}_i \times \text{ Fan bias}_j + \epsilon_{i,j} \tag{11}$$

As for the internal side analysis, we estimate Equation 11 on the entire sample as well as disaggregated by draft round. This enables us to test Prediction 2 from the theoretical framework. We also disaggregate between rounds 1-3 and rounds 4+ because we expect from the theoretical framework to be stronger for early rounds of the draft (Prediction 3). We then compute the share of African-American players drafted by teams and estimate its correlation with the level of fan bias using a linear framework. We then perform the same analysis using our alternative measure of fan bias, the link between the share of African-American players and stadium attendance.

5.4 Cost of racial bias

To evaluate the cost associated with expressing racial bias, we use the WAR (Win Above Replacement) metric, a standard performance measure in professional baseball. Given that many players never reach the major leagues and thus do not have a WAR metric, we assign these players a value of 0. In the appendix, we also conduct the analysis on players who did reach the majors and find similar results.

We estimate the following linear regression for player i drafted by team j:

$$\begin{aligned} WAR_i &= \beta_0 + \beta_R \text{ Race}_i + \gamma \text{ Fan bias}_j + \delta \text{ White}_i \times \text{ Fan bias}_j + \eta \text{ Same race}_{i,j} \\ &\quad + \kappa \text{ ML quality prediction}_i + \epsilon_{i,j} \end{aligned} \tag{12}$$

where Same race indicates whether the player shares the same race as the scouting director. The main coefficients of interest are δ, which measures whether White players drafted by teams with biased fans are of lower quality (if δ is negative), and η, which measures if players drafted by a team whose scouting director shares their race are of lower quality.

6 Empirical Results

Our empirical estimates unfold in a structured four-step process. First, we scrutinize the extensive margin of the draft, probing into whether minority players experience lower probabilities of being drafted compared to their White peers. Second, we examine internal biases within MLB scouting practices. Third, we shift our focus to external biases, assessing how fan biases might influence draft strategies. Finally, we explore the financial implications of these various biases. By dissecting our analysis into these four stages, we provide a thorough overview of the factors impacting player selection and reveal the deep-seated biases within the draft process.

6.1 Extensive margin analysis: probability of reaching the MLB and signing bonus

In the initial phase of our analysis, we focus on the extensive margin, particularly investigating whether minority players have lower chances of being drafted into Major League Baseball (MLB) compared to their White counterparts. Our findings reveal two key results.

Result 1a: There is no statistical difference between races in the probability of being drafted, once we con-

Table 1: Extensive margin analysis

	Drafted (1)
Asian	-0.109
	(0.139)
Black	0.073
	(0.051)
Hispanic	-0.081
	(0.057)
Other races	-0.020
	(0.111)
PG Grade	1.803^{***}
	(0.026)
Constant	-18.271^{***}
	(0.237)
N	61,362
Pseudo-R^2	0.476

Note: The dependent variable is a dummy variable for whether the player was drafted. The variable PG Grade corresponds to the evaluation of players on the Perfect Game website. Racial identity was obtained from players' last names. * p<0.10, ** p<0.05, *** p<0.01. Standard errors in parentheses.

trol for player quality.

Evidence for this result is drawn from the estimation of Equation 7, as shown in Table 1. The data indicate that, when controlling for player quality, there is no correlation between race and the likelihood of being drafted. In fact, if there is any notable difference, it is that African-American players may be slightly more likely than White players to be drafted into the MLB. Bias can, however, manifest in other ways. One notable example is the signing bonuses awarded to players. Here, we uncover a second result.

Result 1b: Controlling for prospect quality, there is no statistical difference in signing bonus amounts based on race.

Using data on signing bonuses, conditional on being drafted, we present findings in Table 2 that show, when controlling for player quality, no racial group receives significantly different signing bonuses compared to White players. Interestingly, controlling for player quality, Asian and White players tend to receive the lowest signing bonuses relative to all other races, as the point estimates for other races are positive.

These findings underscore that, conditional on quality, race does not significantly impact the probability of being drafted or the financial rewards associated with being drafted into MLB. We now shift our focus to analyzing bias from the perspective of teams, starting with managerial bias, specifically biases that favor scouts' own racial groups.

Table 2: Analysis of signing bonuses

	Log(Signing bonus)
	(1)
2 or more races	0.321
	(0.196)
Asian	−0.660
	(0.434)
Black	0.112
	(0.118)
Hispanic	0.183
	(0.137)
Native American	0.210
	(0.368)
ML quality prediction	7.244***
	(0.256)
Constant	10.659***
	(0.079)
N	1,162
R^2	0.415

Note: The dependent variable is the log of the signing bonus obtained by players. The omitted category is White players. The variable "ML quality prediction" is the output of a Machine Learning algorithm computing the probability of a prospect to reach the MLB, as a function of their high-school / college performance. * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses.

6.2 Managerial Bias: Bias in Scouting

In this section, we present our findings on potential racial biases within teams' scouting departments. Our analysis unfolds across two primary areas. First, we explore racial concordance in drafting by investigating whether players are more likely to be drafted by teams whose scouting directors share the same racial background. This step aims to uncover any patterns of racial preference in drafting decisions at the director level. Second, we assess scout evaluations and potential racial bias by examining if individual scouts exhibit favorable evaluations towards players of their own race. This aspect of our analysis seeks to understand the presence of bias at the scout level, which could influence overall team decisions. Through these steps, we strive to provide a comprehensive understanding of racial dynamics within scouting processes and their impact on player selection.

6.2.1 Bias in favor of the scouting director's racial group

First, we examine the draft strategies employed by teams, particularly exploring the influence of the scouting director's race. To perform this analysis, we compiled a synthetic dataset for each team, covering all players considered between 2008 and 2019. Using this dataset, we estimated a logistic regression model to evaluate the probability of a specific team drafting a player. This probability is analyzed while controlling for player quality and a dummy variable that indicates whether the player and the scouting director share the same racial background. Additionally, we account for the player's race and the geographic proximity between the player, the team, and the scouting directors, given the established correlation between these factors and the likelihood of being drafted (see Ahmadi et al. (2022)).

This approach enables us to dissect the nuanced effects of racial concordance on drafting decisions within MLB teams. In doing so, we uncover the next result.

Result 2: Players are more likely to be drafted by teams whose scouting director shares their racial background. This tendency is most pronounced in the later rounds of the draft (rounds 26+).

Evidence for this result is presented in Table 3 (column 1), which indicates a bias in favor of players who share the same race as the scouting director. Further examination in columns 2 to 3 of Table 3 reveals that this racial bias is most apparent in the later rounds of the draft (rounds 26+). This aligns with the findings from (Ahmadi et al., 2022), which demonstrated that propinquity bias—favoring players who are geographically close— plays a more significant role in the lower draft picks, where scouting directors have greater decision-making authority.

When disaggregating the data by time in Table C.1, the bias holds for earlier seasons (2008-2015), albeit without statistical significance at conventional levels, but diminishes in later years. This result aligns with the hypothesis that the impact of scouting directors' racial biases diminishes as advanced statistical models become more prevalent in decision-making. Overall, these findings indicate that scouting directors tend to favor players of their own race, particularly when they hold significant decision-making power over less prominent players. This observation supports Prediction 1 from our theoretical framework, emphasizing the nuanced impact of racial bias in scouting practices.

Table 3: Link between being the same race as the scouting director and the probability of being drafted by a given team

	All rounds	Rounds 1-25	Rounds 26+
	(1)	(2)	(3)
Same race as scouting director	0.050	0.018	0.374*
	(0.076)	(0.081)	(0.226)
2 or more races	0.044	0.016	0.327
	(0.149)	(0.152)	(1.044)
Asian	0.042	0.015	0.313
	(0.329)	(0.347)	(1.040)
African American	0.041	0.015	0.307
	(0.089)	(0.096)	(0.240)
Hispanic	0.041	0.015	0.303
	(0.105)	(0.115)	(0.284)
Native American	0.044	0.016	0.332
	(0.293)	(0.305)	(1.094)
Constant	−3.411***	−3.383***	−3.693***
	(0.090)	(0.096)	(0.291)
Performance metrics and distance controls	Yes	Yes	Yes
N	58,740	51,060	7,470
Pseudo R^2	0.000	0.000	0.001
Mean dep. variable	0.033	0.033	0.033
p-value equality of coefficients		0.13	

Note: The dependent variable is the probability of being drafted, conditional on player quality. The omitted category is White players. * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses. The p-value presented at the bottom of the table corresponds to the p-value of equality of coefficients for the "Same race" variable.

6.2.2 Bias in scouting evaluations

In the next phase of our analysis, we shift our focus to the internal scouting evaluations conducted by a partner Major League Baseball (MLB) team. Specifically, we use the group rating as our outcome variable. This rating system categorizes players on a scale from 1 to 7, where 1 represents the most skilled players and 7 represents the least skilled.

This part of our study seeks to uncover potential biases in the evaluations assigned by individual scouts, specifically examining whether scouts rate players of their own race more favorably than others. By analyzing these internal evaluations, we aim to gain deeper insights into the presence and extent of racial biases within the scouting processes of MLB teams. Upon doing so, we uncover Result 3.

Result 3: Controlling for prospect quality, scouts tend to rate players of their own race more favorably.

Evidence for Result 3 can be found in Table 4. In this regression model, even after controlling for predicted player quality (which is highly correlated with the player's rating), we still observe a significant bias in favor of players who share the same race as the scout (a negative coefficient implies a higher rating). Quantitatively, being the same race as the scout is associated with a 1/5 category higher evaluation. Additionally, African-American and Hispanic scouts are frequently assigned to evaluate lower-quality players. These factors collectively contribute to a strong bias in favor of White players (Table C.2).

Disaggregating the data by draft rounds, as shown in (Table C.3), we find that the scout player concordance effect (bias in favor of same race players) is more pronounced for players selected in rounds 6 and beyond (columns 3 and 4). This finding suggests that scouts' racial bias is stronger for players whose quality is less exceptional, or those who do not stand out as future stars. When disaggregated by time (Table C.4), the data indicate that the bias of scouts in favor of players of their own race has increased over the period studied.

Interestingly, the analysis also indicates that Hispanic, African-American, and mixed-race players tend to receive a "premium" in group ratings, relative to White players, when controlling for prospect quality.

Table 4: Link between race and internal scouting evaluation

	Evaluation (1)
Same race as scout	-0.189^{**}
	(0.074)
ML quality prediction	-5.181^{***}
	(0.172)
Player race	
2 or more races	-0.193
	(0.160)
Asian	-0.206
	(0.380)
Black	-0.154^{*}
	(0.086)
Hispanic	-0.527^{***}
	(0.095)
Native American	-0.352
	(0.308)
Scout race	
Black	0.158^{*}
	(0.089)
Hispanic	-0.039
	(0.087)
Constant	4.872^{***}
	(0.085)
N	1,916
R^2	0.335

Note: The dependent variable is the group scoring from the internal reports for a partner MLB team. A lower value means a better evaluation. The omitted categories are White players and White scouts. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

This premium suggests a favorable bias in the evaluation of these minority players.

6.3 Customer Bias: Fan Bias and Draft Strategy

In this section, we explore another potential source of racial bias in baseball: fan reactions. Our approach leverages a naturally occurring event where all MLB teams posted messages of support for the Black Lives Matter movement in June or July 2020. We collected all related comments from Twitter and Facebook. Using the dataset and methodology from this event, we aim to uncover underlying biases in fan reactions following the methodology in Kumar and Pranesh (2021). This approach yields a measure of bias for each team, defined as the average of negative comments. We then created a composite index based on various measures of negative fan reactions, including replies on twitter, comments on Facebook, and the ratio of "angry" to "like" reactions on Facebook.

To further enrich our analysis, we employed an alternative measure of fan bias by examining the relationship between team attendance and the racial composition of the team. This approach offers a nuanced understanding of how fan bias might impact the dynamics within baseball. The detailed metrics of customer bias are provided in Section 4.

6.3.1 Fan bias and Team Draft Strategy

Our initial test examines the relationship between fan bias and team draft strategy. This analysis yields a first finding: Result 4.

Result 4: African-American players are less likely to be drafted by teams with more biased fans. This effect is most evident statistically in the early rounds of the draft(rounds 1-3).

Evidence supporting Result 4 is obtained from the estimation of Equation 11, and summary estimates are displayed in Table 5. Empirical results indicate that African-American players are less likely to be drafted by teams with biased fans. This finding aligns with Prediction 2 of our theoretical framework.

By disaggregating the data between early and later rounds of the draft, we can ascertain whether customer preferences for race and quality function as substitutes or complements (Equation 1). Interestingly, the correlation is stronger in *early* rounds of the draft (Rounds 1-3, Column 2), where fan scrutiny is more intense. This contrasts with the findings for internal bias. In the later rounds of the draft, although the point estimate remains negative, the magnitude is significantly reduced (Column 3). This pattern suggests that fans perceive race and quality of players as complementary factors in their utility function, aligning with Prediction 3.

Table 5: Fan bias and probability of being drafted

	All rounds (1)	Rounds 1-3 (2)	Rounds 4+ (3)
2 or more races	-0.078	-0.119	-0.045
	(0.134)	(0.206)	(0.180)
Asian	0.020	0.028	0.017
	(0.323)	(0.458)	(0.456)
Black	-0.015	-0.065	-0.011
	(0.059)	(0.135)	(0.068)
Hispanic	0.002	-0.024	0.007
	(0.076)	(0.176)	(0.090)
Native American	-0.057	0.030	-0.106
	(0.283)	(0.458)	(0.360)
ML quality prediction	-0.050	-0.009	-0.066
	(0.150)	(0.268)	(0.303)
Index of fan bias	0.029	0.046	0.026
	(0.025)	(0.049)	(0.029)
Black × Index of fan bias	-0.229^{***}	-0.497^{***}	-0.164^{**}
	(0.062)	(0.142)	(0.070)
Constant	-3.344^{***}	-3.351^{***}	-3.342^{***}
	(0.042)	(0.116)	(0.058)
Scouting director race	Yes	Yes	Yes
N	64,757	15,515	46,806
Pseudo R^2	0.001	0.004	0.001
Mean dep. variable	0.033	0.033	0.033
p-value equality of coefficients		0.04	

Note: The dependent variable is a dummy of whether the player was drafted by a given team. The dependent variable Index of fan bias is defined above. The other dependent variables are player's race (White is the omitted category) and prospect quality. * p<0.10, ** p<0.05, *** p<0.01. Standard errors in parentheses. The p-value presented at the bottom of the table corresponds to the p-value of equality of coefficients for the "Black × Index" variable.

6.3.2 Aggregate effect of fan bias

Our data are sufficiently rich to examine the aggregate effect of fan bias. Upon doing so, we report a fifth key result.

Result 5: Overall, teams with more biased fans are less likely to draft African-American players, particularly in the early rounds of the draft.

Building on our previous analysis of fan bias and its impact on draft probabilities, we now present the aggregate findings. We regressed the total share of African-American players drafted by teams on the index of fan bias, derived from reactions to Black Lives Matter posts. Empirical results, summarized in Table 6 and in Figure 1, reveal a clear relationship between fan bias and the racial composition of drafted players.

To sum up, our analysis uncovers a negative correlation between the team bias index and the proportion of African-American players drafted. Specifically, a one standard deviation increase in fan bias is associated with a 3.6 percentage point reduction in the share of African-American players drafted by a team. This effect size accounts for approximately 20% of the league average or 0.41 standard deviations, underscoring the substantial impact of fan bias on draft outcomes.

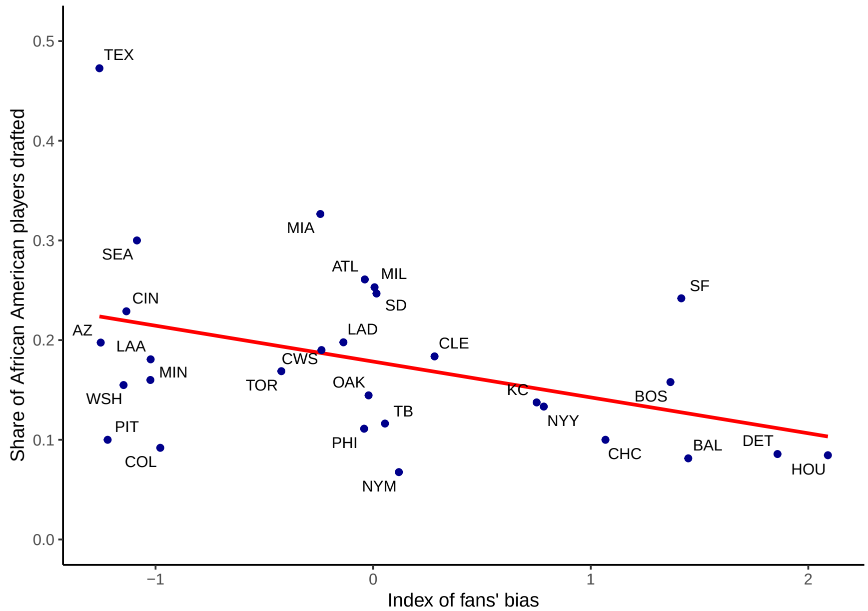
Table 6: Correlation fans bias and draft of Black players by round of draft

	All rounds (1)	Rounds 1-3 (2)	Rounds 4+ (3)
Index	-0.036^{**}	-0.063^{***}	-0.028
	(0.016)	(0.023)	(0.017)
Constant	0.178^{***}	0.155^{***}	0.189^{***}
	(0.015)	(0.022)	(0.017)
N	29	29	29
R^2	0.164	0.224	0.091

Note: The dependent variable in Column 1 is the share of African American players drafted in any round of the draft. In Column 2, it is the share of African American players drafted in the first three rounds. In Column 3, it is the share of African American players drafted in later rounds (4+). The independent variable is the index of fans' bias, defined above. No Facebook post related to Black Lives Matter was found for the St Louis Cardinals, explaining why the sample size is 29. * p<0.10, ** p<0.05, *** p<0.01. Standard errors in parentheses.

After examining the overall relationship between fan bias and team draft strategy, we further dissected

Figure 1: Index of fan bias and share of African-American players drafted



Note: The X-axis corresponds to the composite index of fan bias computed with reactions to posts related to the Black Lives Matter movement. The Y-axis corresponds to the share of African-American players drafted during the study period (2008-20). The red line is the regression line. $\beta = -0.036$, $p = 0.03$.

the relationship to see if it varies between early and later draft picks. The rationale for this differentiation is that players selected in the early rounds of the draft (particularly within the first three rounds) have a higher probability of making it to the MLB. These players attract greater media scrutiny and hold more financial significance for teams.

By examining these distinctions, we aim to discern whether fan bias exerts a more pronounced influence on draft decisions for players likely to become prominent figures in the league, compared to those drafted later who might not attract as much public scrutiny or financial investment.

Empirical results presented in Table 6 highlight a strong correlation between fan bias and the drafting of African-American players in the early rounds (column 2), while this correlation is not significant in later rounds, despite the negative coefficient (column 3). These findings suggest that team owners might consider fan bias when drafting highly publicized players. The robustness of these results for early rounds, even when controlling for fan base composition and local characteristics (Table D.1), further underscores this point.

At this point, it is crucial to note that we cannot definitively establish causality. The observed correlation could suggest that biased fans influence drafting decisions. However, it is also possible that fans develop biases due to the underrepresentation of African-American players on their favorite teams.

6.3.3 Robustness check using fan bias in attendance

While the empirical results presented thus far suggest that external bias driven by fans is associated with draft decisions, a critic might question whether our findings are influenced by our specific measure of bias. To address this concern, we developed an alternative measure of fan bias—attendance. Upon doing so, we find Result 6.

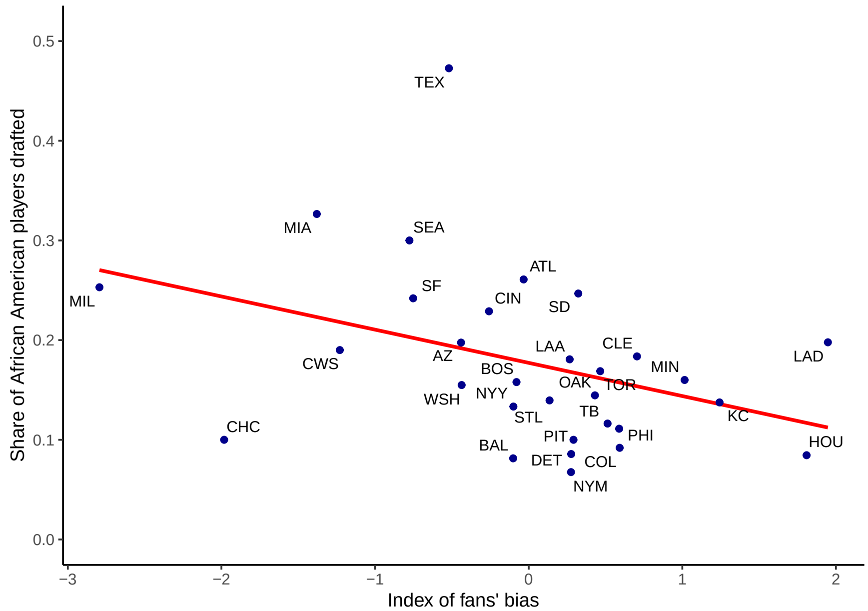
Result 6: As a whole, the data suggest that teams with more biased fans, as measured by stadium attendance, draft fewer African-American players

To provide evidence for this result, we complement our previous analysis of the link between fan bias and drafting strategies by performing similar estimations. This time, we use an alternative measure of fan bias based on stadium attendance. The measure, defined in Section 4, allows us to further examine the relationship between fan bias and draft outcomes. For comparability, the coefficient is normalized and inverted, so that a positive coefficient corresponds to more biased fans.[6]

We find that the attendance bias is strongly correlated with the share of African-American players drafted, even after controlling for local factors (Figure 2, Table E.2, columns 1 and 2). Interestingly, both index measures remain correlated with the share of African-American players drafted (Table E.2, column 3). Disaggregating by draft round (Table E.3), we do not find statistically significant differences between the correlation coefficients. In conclusion, utilizing an alternative measure of fan bias confirms the link between fan bias and draft strategies. Perhaps not surprisingly, but the data suggest that team executives may indeed consider fan preferences when making draft decisions.

[6]Our two measures of fan bias (the index from the BLM posts) and the bias in attendance are positively correlated, although insignificantly (Table E.1).

Figure 2: Attendance bias and share of African-American players drafted



Note: The X-axis corresponds to the attendance bias (definition in Section 4). The Y-axis corresponds to the share of African-American players drafted during the study period (2008-20). The red line is the regression line. $\beta = -0.033$, $p = 0.04$.

6.4 Cost of racial bias

In our final analysis, we aim to quantify the economic cost of the observed racial bias. We achieve this by estimating the value players contribute to their team's performance using WAR (Wins Above Replacement) as the metric for their first year in the majors. By analyzing these performance metrics, we provide a comprehensive understanding of the financial impact associated with racial bias in player drafting decisions. Two key results emerge from this analysis.

Result 7a: Teams with biased fans draft less qualified White players, leading to compromised competitiveness.

Result 7b: Homophily of the scouting director does not result in drafting less qualified players. The bias predominantly impacts those with a low probability of reaching the majors. Although slight quality deviations are observed in the later draft rounds, the sample size is small, limiting the overall impact.

Table 7 summarizes the key estimation results. As can be seen in Column 1 of Table 7, White players drafted by teams with more biased fans tend to under-perform compared to those selected by less biased teams. Conversely, non-White players drafted by teams with more biased fans tend to perform better, though the results are only marginally significant ($p = 0.09$ for both variables).

In terms of magnitude, a one standard deviation increase in fan bias for white players is associated with an approximate 0.2 WAR decrease. Considering a conservative value of $5 million per WAR (Steinberg, 2022), this estimate indicates that teams are effectively willing to draft players who are worth about $1 million less as fan bias increases by one standard deviation.

Our analysis of internal bias, as detailed in Table 7, shows no correlation between players' WAR and their race relative to the scouting director. In the later rounds of the draft (column 4), where internal bias was noted, we observe a negative coefficient, suggesting that players of the same race as the scouting director perform slightly worse. However, the sample size is too small for definitive conclusions.

In summary, teams with biased fans are willing to incur significant financial costs by drafting less qualified White players, while internal bias (in favor of the scouting director's race) does not seem to impact financial outcomes significantly. For rational, profit-maximizing teams, this implies that fan bias must compensate for the lost prospect quality in terms of increased profits when drafting White players.

Table 7: Cost of racial bias in performance of drafted players

	All rounds (1)	Rounds 1-5 (2)	Rounds 6-15 (3)	Rounds 16+ (4)
Index of fan bias	0.172*	0.162	0.150	0.520
	(0.103)	(0.145)	(0.191)	(0.845)
White players × Index of fan bias	−0.204*	−0.212	−0.210	−0.328
	(0.118)	(0.175)	(0.203)	(0.886)
Same race as scouting director	0.093	−0.106	0.267	−0.974**
	(0.177)	(0.305)	(0.291)	(0.459)
Constant	0.202	0.454	−0.178	1.567**
	(0.187)	(0.323)	(0.321)	(0.631)
Race-fixed effects	Yes	Yes	Yes	Yes
Prospect quality metric	Yes	Yes	Yes	Yes
N	323	168	83	28
R^2	0.049	0.076	0.102	0.352

Note: The dependent variable is the WAR (Win Above Replacement) value. The index of fan bias is defined in Section 4. Player's race-fixed effects are included (White is the omitted category). * p<0.10, ** p<0.05, *** p<0.01. Standard errors in parentheses.

7 Conclusion

In this study, we empirically test two channels of racial bias in the context of the Major League Baseball draft using a set of publicly available and proprietary data. Several insights emerge. First, our finding that customer bias against African Americans influences drafting decisions in the early rounds of the MLB draft underscores the significant impact of fan preferences on team strategies. This reveals how deeply entrenched biases can affect organizational decisions, even at the cost of team performance. Importantly, this bias has economic implications. For example, the fact that teams with more biased fans are willing to draft lower-quality white players at a substantial cost highlights a critical economic dimension. This result shows that catering to biased customer preferences can lead to inefficient resource allocation, ultimately impacting the team's competitive edge.

Second, we argue that our results deepen our understanding of homophily: the observation that scouting directors' homophily occurs in later rounds, where scrutiny is lower, offers insights into how internal biases manifest when external pressures are reduced. This indicates that biases within the organization can influence decisions when there is less public oversight, which is an important consideration for addressing discriminatory practices. These results also further indicate that firms' decisions respond to incentives, as internal biases (which are inefficient for the firms' profits) only occur when scrutiny and financial consequences are low.

Third, we view our results as informing policy makers and sports management about the need for interventions to mitigate bias in drafting decisions. Understanding the contexts and circumstances in which biases influence decisions allows for the development of more effective strategies to promote fairness and equity.

Even though our results are drawn from a specific corner of the world, we view our results as having import in several different areas. For example, our theoretical and empirical combination can provide a framework for beginning an analysis of many important economic settings where multiple sources of bias interact simultaneously. A key feature is that our framework reveals that examining bias at the aggregate level may mask different and potentially conflicting sources of bias. We also find our results as plausibly generalizing to other markets. Following List (2020) in considering the generalizability of our insights, we view preferences, beliefs, and constraints within our market as similarly situated to other industries, whether business to business or business to customer. The fundamental nature of the trade-offs that we document are shared broadly. In this manner, a general lesson learned is that scholars who examine data for discrimination using establishment level data run a certain risk when measuring aggregates. If there is a tension in biased preferences between management and customers, then key aggregates can underestimate, or mask, key biases.

References

Aggarwal, P., A. Brandon, A. Goldszmidt, J. Holz, J. A. List, I. Muir, G. Sun, and T. Yu (2022). Using high-frequency location data to evaluate racial bias in policing. Technical report, NBER Working Paper.

Ahmadi, M., N. Durst, J. Lachman, J. A. List, M. List, N. List, and A. T. Vayalinkal (2022). Nothing propinks like propinquity: Using machine learning to estimate the effects of spatial proximity in the major league baseball draft. Technical report, National Bureau of Economic Research.

Akerlof, G. A. (1997). Social distance and social decisions. *Econometrica: Journal of the Econometric Society 65*(5), 1005–1027.

Arrow, K. J. (1972). Models of job discrimination. In *Racial discrimination in economic life*. Lexington, Massachusetts.

Auspurg, K., A. Schneck, and T. Hinz (2019). Closed doors everywhere? a meta-analysis of field experiments on ethnic discrimination in rental housing markets. *Journal of Ethnic and Migration Studies 45*(1), 95–114.

Avery, D. R., P. F. McKay, S. Tonidandel, S. D. Volpone, and M. A. Morris (2012). Is there method to the madness? examining how racioethnic matching influences retail store productivity. *Personnel Psychology 65*(1), 167–199.

Bandiera, O., I. Barankay, and I. Rasul (2009). Social connections and incentives in the workplace: Evidence from personnel data. *Econometrica 77*(4), 1047–1094.

Bar, R. and A. Zussman (2017). Customer discrimination: evidence from israel. *Journal of Labor Economics 35*(4), 1031–1059.

Becker, G. S. (1971). *The economics of discrimination*. University of Chicago press.

Bell, R. R. (1981). Friendships of women and of men. *Psychology of Women Quarterly 5*(3), 402–417.

Benson, A., S. Board, and M. Meyer-ter Vehn (2024). Discrimination in hiring: Evidence from retail sales. *Review of Economic Studies 91*(4), 1956–1987.

Bohren, J. A., P. Hull, and A. Imas (2022). Systemic discrimination: Theory and measurement. Technical report, National Bureau of Economic Research.

Bohren, J. A., A. Imas, and M. Rosenberg (2019). The dynamics of discrimination: Theory and evidence. *American economic review 109*(10), 3395–3436.

Bourabain, D. and P.-P. Verhaeghe (2019). Could you help me, please? intersectional field experiments on everyday discrimination in clothing stores. *Journal of Ethnic and Migration Studies 45*(11), 2026–2044.

Boyd-Swan, C. and C. M. Herbst (2019). Racial and ethnic discrimination in the labor market for child care teachers. *Educational Researcher 48*(7), 394–406.

Brown, E., R. Spiro, and D. Keenan (1991). Wage and nonwage discrimination in professional basketball: do fans affect it? *American Journal of Economics and Sociology 50*(3), 333–345.

Brown, R. W. and R. T. Jewell (1995). Race, revenues, and college basketball. *The Review of Black Political Economy 23*(3), 75–90.

Broyles, P. and B. Keen (2010). Consumer discrimination in the nba: An examination of the effect of race on the value of basketball trading cards. *The Social Science Journal 47*(1), 162–171.

Bryson, A. and A. Chevalier (2015). Is there a taste for racial discrimination amongst employers? *Labour Economics 34*, 51–63.

Burdekin, R. C., R. T. Hossfeld, and J. K. Smith (2005). Are nba fans becoming indifferent to race? evidence from the 1990s. *Journal of sports economics 6*(2), 144–159.

Burdekin, R. C. and T. L. Idson (1991). Customer preferences, attendance and the racial structure of professional basketball teams. *Applied Economics 23*(1), 179–186.

Bygren, M. (2010). The gender composition of workplaces and men's and women's turnover. *European Sociological Review 26*(2), 193–202.

Combes, P.-P., B. Decreuse, M. Laouenan, and A. Trannoy (2016). Customer discrimination and employment outcomes: theory and evidence from the french labor market. *Journal of Labor Economics 34*(1), 107–160.

Day, D. V., S. Gordon, and C. Fink (2012). The sporting life: Exploring organizations through the lens of sport. *Academy of Management Annals 6*(1), 397–433.

Depken II, C. A. and J. M. Ford (2006). Customer-based discrimination against major league baseball players: Additional evidence from all-star ballots. *The Journal of Socio-Economics 35*(6), 1061–1077.

Dey, M. S. (1997). Racial differences in national basketball association players' salaries: A new look. *The American Economist 41*(2), 84–90.

Doleac, J. L. and L. C. Stein (2013). The visible hand: Race and online market outcomes. *The Economic Journal 123*(572), F469–F492.

FitzGerald, C. and S. Hurst (2017). Implicit bias in healthcare professionals: a systematic review. *BMC medical ethics 18*(1), 1–18.

Fonti, F., J.-M. Ross, and P. Aversa (2023). Using sports data to advance management research: A review and a guide for future studies. *Journal of Management 49*(1), 325–362.

Giuliano, L., D. I. Levine, and J. Leonard (2009). Manager race and the race of new hires. *Journal of Labor Economics 27*(4), 589–631.

Glennon, B., F. Morales, S. Carnahan, and E. Hernandez (2024). Does employing skilled immigrants enhance competitive performance? evidence from european football clubs. *Management Science*.

Glover, D., A. Pallais, and W. Pariente (2017). Discrimination as a self-fulfilling prophecy: Evidence from french grocery stores. *The Quarterly Journal of Economics 132*(3), 1219–1260.

Goldin, C. and C. Rouse (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American economic review 90*(4), 715–741.

Gomez-Gonzalez, C., C. Nesseler, and H. M. Dietl (2021). Mapping discrimination in europe through a field experiment in amateur sport. *Humanities and Social Sciences Communications 8*(1), 1–8.

Hamilton, B. H., J. A. Nickerson, and H. Owan (2012). Diversity and productivity in production teams. In *Advances in the Economic Analysis of participatory and Labor-managed Firms*, pp. 99–138. Emerald Group Publishing Limited.

Hauptman, D. (2009). The need for a worldwide draft to level the playing field and strike out the national origin discrimination in major league baseball. *Loy. LA Ent. L. Rev. 30*, 263.

Hedegaard, M. S. and J.-R. Tyran (2018). The price of prejudice. *American Economic Journal: Applied Economics 10*(1), 40–63.

Hjort, J. (2014). Ethnic divisions and production in firms. *The Quarterly Journal of Economics 129*(4), 1899–1946.

Hoang, H. and D. Rascher (1999). The nba, exit discrimination, and career earnings. *Industrial Relations: A Journal of Economy and Society 38*(1), 69–91.

Holzer, H. J. and K. R. Ihlanfeldt (1998). Customer discrimination and employment outcomes for minority workers. *The Quarterly Journal of Economics 113*(3), 835–867.

Jowell, R. and P. Prescott-Clarke (1970). Racial discrimination and white-collar workers in britain. *Race 11*(4), 397–417.

Kahn, L. M. (2000). The sports business as a labor market laboratory. *Journal of economic perspectives 14*(3), 75–94.

Kahn, L. M. (2009). The economics of discrimination: Evidence from basketball. Technical report, IZA Discussion Paper.

Kahn, L. M. and M. Shah (2005). Race, compensation and contract length in the nba: 2001–2002. *Industrial Relations: A Journal of Economy and Society 44*(3), 444–462.

Kahn, L. M. and P. D. Sherer (1988). Racial differences in professional basketball players' compensation. *Journal of Labor Economics 6*(1), 40–61.

Kanazawa, M. T. and J. P. Funk (2001). Racial discrimination in professional basketball: Evidence from nielsen ratings. *Economic Inquiry 39*(4), 599–608.

Kelley, E. M., G. V. Lane, M. Pecenco, and E. A. Rubin (2024). Customer discrimination in the workplace: Evidence from online sales. *Journal of Labor Economics 5*(42), –.

Kumar, S. and R. R. Pranesh (2021). Tweetblm: A hate speech dataset and analysis of black lives matter-related microblogs on twitter. Technical report, arXiv preprint arXiv:2108.12521.

Kuppuswamy, V. and P. Younkin (2020). Testing the theory of consumer discrimination as an explanation for the lack of minority hiring in hollywood films. *Management Science 66*(3), 1227–1247.

Landry, C. E., A. Lange, J. A. List, M. K. Price, and N. G. Rupp (2006). Toward an understanding of the economics of charity: Evidence from a field experiment. *The Quarterly journal of economics 121*(2), 747–782.

Lang, K. and A. K.-L. Spitzer (2020). Race discrimination: An economic perspective. *Journal of Economic Perspectives 34*(2), 68–89.

Laouenan, M. (2017). ‘hate at first sight’: Evidence of consumer discrimination against african-americans in the us. *Labour Economics 46*, 94–109.

Leonard, J. S., D. I. Levine, and L. Giuliano (2010). Customer discrimination. *The Review of Economics and Statistics 92*(3), 670–678.

Liebe, U. and H. Beyer (2021). Examining discrimination in everyday life: A stated choice experiment on racism in the sharing economy. *Journal of Ethnic and Migration Studies 47*(9), 2065–2088.

Lippens, L., S. Baert, and E. Derous (2021). Loss aversion in taste-based employee discrimination: Evidence from a choice experiment. *Economics Letters 208*, 110081.

List, J. A. (2004). The nature and extent of discrimination in the marketplace: Evidence from the field. *The Quarterly Journal of Economics 119*(1), 49–89.

List, J. A. (2020). Non est disputandum de generalizability? a glimpse into the external validity trial. Technical report, National Bureau of Economic Research.

List, J. A. (2025). *Experimental Economics: Theory and Practice*. University of Chicago Press.

MacLeod, S. A. and P. W. Newall (2022). Investigating racial bias within australian rules football commentary. *PLoS one 17*(7), e0272005.

Maennig, W. and S. Q. Mueller (2021). Consumer and employer discrimination in professional sports markets–new evidence from major league baseball. Technical Report 69, Hamburg Contemporary Economic Discussions.

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

McLoughlin, D. (2020). Racial bias in football commentary (study).

McPherson, M., L. Smith-Lovin, and J. M. Cook (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology 27*(1), 415–444.

Nardinelli, C. and C. Simon (1990). Customer racial discrimination in the market for memorabilia: The case of baseball. *The Quarterly Journal of Economics 105*(3), 575–595.

Parrett, M. (2011). Customer discrimination in restaurants: Dining frequency matters. *Journal of Labor Research 32*(2), 87–112.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The american economic review 62*(4), 659–661.

Potârcă, G. and M. Mills (2015). Racial preferences in online dating across european countries. *European Sociological Review 31*(3), 326–341.

Principe, F. and J. C. van Ours (2022). Racial bias in newspaper ratings of professional football players. *European Economic Review 141*, 103980.

Quillian, L., D. Pager, O. Hexel, and A. H. Midtbøen (2017). Meta-analysis of field experiments shows no change in racial discrimination in hiring over time. *Proceedings of the National Academy of Sciences 114*(41), 10870–10875.

Riach, P. A. and J. Rich (2002). Field experiments of discrimination in the market place. *The economic journal 112*(483), F480–F518.

Rosenman, E. T., S. Olivella, and K. Imai (2022). Race and ethnicity data for first, middle, and last names.

Ross, S. F. and M. James Jr (2015). A strategic legal challenge to the unforeseen anticompetitive and racially discriminatory effects of baseball's north american draft. *Colum. L. Rev. Sidebar 115*, 127.

Schmidt, A. and K. Coe (2014). Old and new forms of racial bias in mediated sports commentary: The case of the national football league draft. *Journal of Broadcasting & Electronic Media 58*(4), 655–670.

Silverman, A. (2020). Demographic data shows which major sports fan bases are most likely to support or reject social justice advocacy. Technical report, Morning Consult.

Steinberg, R. (2022). Mlb has lowest percentage of african american baseball players in 3 decades: Report.

Sur, P. K. and M. Sasaki (2020). Measuring customer discrimination: evidence from the professional cricket league in india. *Journal of Sports Economics 21*(4), 420–448.

Zussman, A. (2013). Ethnic discrimination: Lessons from the israeli online market for used cars. *The Economic Journal 123*(572), F433–F468.

Appendices

A Descriptive statistics

Table A.1: Descriptive statistics

Variable	Mean	Std. Dev.	N
Panel A. Player race			
White	0,696	0,460	2233
Asian	0,004	0,067	2233
Black	0,171	0,376	2233
Hispanic	0,093	0,290	2233
Native American	0,006	0,079	2233
2 or more races	0,030	0,169	2233
Panel B. Metrics of player quality			
ML predicted quality	0,206	0,147	2233
Group evaluation	3,555	1,377	1912
Panel C. Signing bonus			
Signing bonus	567 691,04	913 114,63	2153
Panel D. Rounds of the draft			
Draft round	11,928	10,178	2149
Rounds 1-3	0,249	0,433	2149
Rounds 26+	0,133	0,339	2149

B Definitions of fan bias

Table B.1: Examples of comments with bias metric

Text	**Bias**
Love this team.	0
Over 600 likes better than I expected.	0
SO I guess the Blue Jays are a bunch of communists then? I sure hope lots of leftists watch baseball haha	1
I support the one player that did not join the BLM sheep.	1

Note: The "Bias" measure takes the value 1 if the algorithm estimates that the answer is biased. The estimation is done using Kumar and Pranesh (2021).

Table B.2: Restricting to the replies of followers

	Bias replies tweets (1)
Bias only followers	0.517***
	(0.088)
Constant	0.120***
	(0.021)
N	30
R^2	0.552

Note: The dependent variable is the average level of bias for replies to the tweets of teams involving Black Lives Matter. The independent variable is the same share, but restricted to comments posted by people following the team. * p<0.10, ** p<0.05, *** p<0.01. Standard errors in parentheses.

Table B.3: Index of fan bias

Team	Index	Team	Index	Team	Index
Arizona Diamondbacks	-1,252	Houston Astros	2,090	Philadelphia Phillies	-0,041
Atlanta Braves	-0,038	Kansas City Royals	0,752	Pittsburgh Pirates	-1,220
Baltimore Orioles	1,449	Los Angeles Angels	-1,023	San Diego Padres	0,016
Boston Red Sox	1,367	Los Angeles Dodgers	-0,137	San Francisco Giants	1,417
Chicago Cubs	1,068	Miami Marlins	-0,243	Seattle Mariners	-1,086
Chicago White Sox	-0,237	Milwaukee Brewers	0,007	St. Louis Cardinals	
Cincinnati Reds	-1,134	Minnesota Twins	-1,024	Tampa Bay Rays	0,054
Cleveland Guardians	0,282	New York Mets	0,118	Texas Rangers	-1,258
Colorado Rockies	-0,978	New York Yankees	0,784	Toronto Blue Jays	-0,422
Detroit Tigers	1,859	Oakland Athletics	-0,021	Washington Nationals	-1,147

Note: The index of fan bias is defined using the reactions to the posts related to Black Lives Matter, as defined in Section 6.3. A higher coefficient means that fans reacted more negatively to the posts.

Table B.4: Fan bias in attendance

Team	Bias	Team	Bias	Team	Bias
Arizona Diamondbacks	-0.440	Houston Astros	1.809	Philadelphia Phillies	0.589
Atlanta Braves	-0.033	Kansas City Royals	1.242	Pittsburgh Pirates	0.292
Baltimore Orioles	-0.101	Los Angeles Angels	0.267	San Diego Padres	0.323
Boston Red Sox	-0.079	Los Angeles Dodgers	1.948	San Francisco Giants	-0.752
Chicago Cubs	-1.982	Miami Marlins	-1.379	Seattle Mariners	-0.776
Chicago White Sox	-1.229	Milwaukee Brewers	-2.794	St. Louis Cardinals	0.136
Cincinnati Reds	-0.258	Minnesota Twins	1.015	Tampa Bay Rays	0.514
Cleveland Guardians	0.705	New York Mets	0.276	Texas Rangers	-0.519
Colorado Rockies	0.592	New York Yankees	-0.099	Toronto Blue Jays	0.466
Detroit Tigers	0.277	Oakland Athletics	0.431	Washington Nationals	-0.436

Note: The attendance bias is defined as the normalized, inverted coefficient of the regression of the share of African American players on attendance. See description in Section 6.3. A higher coefficient means that when the share of African American players increase, fewer people attend games.

C Internal Bias: Additional Results

Table C.1: Same race as the scouting director and probability of being drafted, by years

	All rounds (1)	Rounds 1-25 (2)	Rounds 26+ (3)
Panel A. Baseline: All years combined			
Same race as scouting director	0.050	0.018	0.374*
	(0.076)	(0.081)	(0.226)
Panel B. Early seasons (2008-2014)			
Same race as scouting director	0.019	0.017	1.122
	(0.190)	(0.197)	(0.691)
Panel C. Late seasons (2015-2020)			
Same race as scouting director	0.057	0.019	0.299
	(0.084)	(0.089)	(0.233)

Note: The outcome variable is the probability of being drafted by each team. Controls include player race fixed effects, distance to the team and the ML prediction of player quality. Panel A represents the baseline estimation, with all years included. Panels B and C represent the results from similar estimations, with players divided with a median split with respect to the season they are drafted. In Panel B, all players drafted before the season 2015 are used. For Panel C, all players drafted after the season 2015 are used. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Table C.2: Race of scouts and player quality

	ML quality predictor
	(1)
Black scout	-0.028^{***}
	(0.004)
Hispanic scout	-0.012^{***}
	(0.004)
Constant	0.171^{***}
	(0.001)
N	8,553
R^2	0.005

Note: The dependent variable is the predicted quality of the prospect, coming from the ML algorithm taking as output high school or college data and trained on the players who made it to the MLB. A higher value of the outcome is associated with a higher quality of prospect. The omitted category is White scouts. * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses.

Table C.3: Same race as scouts, by draft round

	All rounds	Rounds 1-5	Rounds 6-15	Rounds 16+
	(1)	(2)	(3)	(4)
Same race as scout	-0.189^{**}	-0.131	-0.152	-0.182
	(0.074)	(0.123)	(0.132)	(0.139)
ML quality prediction	-5.181^{***}	-3.057^{***}	-4.350^{***}	-4.391^{***}
	(0.172)	(0.204)	(0.643)	(0.676)
Player race				
2 or more races	-0.193	-0.103	0.151	0.116
	(0.160)	(0.187)	(0.581)	(0.587)
Asian	-0.206	-0.415	0.075	0.038
	(0.380)	(0.374)	(1.138)	(1.149)
Black	-0.154^{*}	-0.127	-0.512^{***}	-0.495^{***}
	(0.086)	(0.141)	(0.147)	(0.156)
Hispanic	-0.527^{***}	-0.326^{**}	-0.679^{***}	-0.653^{***}
	(0.095)	(0.146)	(0.174)	(0.184)
Native American	-0.352	0.076	-0.384	0.257
	(0.308)	(0.326)	(0.810)	(1.149)
Scout race				
Black scout	0.158^{*}	-0.077	0.119	0.109
	(0.089)	(0.146)	(0.159)	(0.168)
Hispanic scout	-0.039	0.113	-0.060	-0.138
	(0.087)	(0.129)	(0.161)	(0.171)
Constant	4.872^{***}	3.680^{***}	5.295^{***}	5.336^{***}
	(0.085)	(0.141)	(0.155)	(0.162)
N	1,916	683	508	450
R^2	0.335	0.257	0.121	0.125

Note: The dependent variable is the group scoring from the internal reports for a partner MLB team. A lower value means a better evaluation. The omitted category is White players. Each column corresponds to different subsamples with respect to the drafting round. * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses.

Table C.4: Relationship between race and scouting evaluation, evolution with time

	Baseline (All seasons) (1)	Early seasons (2008-2014) (2)	Late seasons (2015-2020) (3)
Same race as scout	-0.254^{***}	-0.154^{**}	-0.253^{***}
	(0.054)	(0.075)	(0.071)
2 or more races	-0.280^{**}	-0.188	-0.232
	(0.137)	(0.159)	(0.202)
Asian	-0.313	0.349	-0.496
	(0.307)	(0.499)	(0.373)
Black	-0.243^{***}	-0.021	-0.255^{***}
	(0.071)	(0.090)	(0.098)
Hispanic	-0.530^{***}	-0.453^{***}	-0.543^{***}
	(0.081)	(0.104)	(0.110)
Native American	-0.402	-0.592	-0.309
	(0.286)	(0.434)	(0.356)
ML quality prediction	-4.592^{***}	-3.058^{***}	-5.292^{***}
	(0.143)	(0.177)	(0.203)
Constant	4.803^{***}	4.038^{***}	5.095^{***}
	(0.059)	(0.085)	(0.076)
N	2,493	913	1,580
R^2	0.301	0.264	0.309

Note: The dependent variable is the group scoring from the internal reports for a partner MLB team. A lower value means a better evaluation. The omitted category is White players. In column 1, all the sample of players is included. In column 2, players drafted before 2015 (the median split of our sample) are used. In column 3, players drafted after 2015 are used. * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses.

D External Bias: Robustness check with the inclusion of controls

To test for the robustness of our findings, we controlled for two types of factors: (1) indicators of racial and political composition of fanbases; (2) indicators of the local situation of teams with the Share of African-Americans in the county of the team and the vote share for Donald Trump in the 2020 presidential election. Results are presented in Table D.1. Including controls of fans' demographic or political variables does not change the influence of fan bias on teams' recruitment of African-American players in the early stages of the draft (column 2). Neither does the inclusion of the share of African-Americans and Republicans at the county level (column 3).

Table D.1: Robustness of the relationship between fan bias and draft decisions

	Dependent variable: Share of African American players drafted		
	Baseline	Fans indicators	Local indicators
	(1)	(2)	(3)
Index of fan bias	-0.036^{**}	-0.035^{**}	-0.033^{*}
	(0.016)	(0.016)	(0.017)
Share Black fans		0.500	
		(0.752)	
Share Republican fans		0.672	
		(0.583)	
Share local Blacks			-0.044
			(0.135)
Share local Republicans			0.089
			(0.157)
Constant	0.178^{***}	-0.118	0.163^{**}
	(0.015)	(0.243)	(0.062)
N	29	29	28
R^2	0.164	0.212	0.188

Note: The dependent variable is the share of African American players drafted in all rounds of the draft. The dependent variable Index is the metric of fans' bias. The other dependent variables are: in Column 2, the shares of African Americans and Republicans among fans; in Column 3, the share of African Americans and Republicans at the county level. * p<0.10, ** p<0.05, *** p<0.01. Standard errors in parentheses.

E External Bias: Robustness check with Attendance Bias

Table E.1: Correlation measures of fan bias

	Bias in attendance
Index	0.096
	(0.195)
Constant	−0.005
	(0.192)
N	29
R^2	0.009

Note: The dependent variable is the bias of fans in attendance, measured as the correlation between attendance and the share of African American players in the team. A higher value of the index means that increasing the share of African American players is associated with lower attendance. The independent variable is the index of bias, measured using a textual analysis of replies to posts related to the Black Lives Matter. * p<0.10, ** p<0.05, *** p<0.01. Standard errors in parentheses.

Table E.2: Attendance bias and draft strategy

	Dependent variable: Share of African American players drafted		
	Attendance bias	Including local controls	Two bias indices
	(1)	(2)	(3)
Attendance bias	-0.033^{**}	-0.037^{**}	-0.030^{**}
	(0.015)	(0.015)	(0.015)
Share local Blacks		-0.130	
		(0.120)	
Share local Republicans		0.128	
		(0.148)	
Index			-0.033^{**}
			(0.015)
Constant	0.177^{***}	0.170^{***}	0.178^{***}
	(0.015)	(0.059)	(0.015)
N	30	29	29
R^2	0.144	0.242	0.281

Note: The dependent variable is the share of African American players drafted in any round of the draft. The dependent variable attendance bias is measured as the correlation between attendance and the share of African American players in the team. The additional independent variable in column 3 is the index of fan bias from reactions to Black Lives Matter posts. * p<0.10, ** p<0.05, *** p<0.01. Standard errors in parentheses.

Table E.3: Attendance bias and draft of Black players by draft round

	All rounds	Rounds ≤ 3	Rounds 4+
	(1)	(2)	(3)
Attendance bias	-0.033^{**}	-0.037	-0.035^{**}
	(0.015)	(0.024)	(0.016)
Constant	0.177^{***}	0.151^{***}	0.189^{***}
	(0.015)	(0.024)	(0.016)
N	30	30	30
R^2	0.144	0.077	0.145

Note: The dependent variable in Column 1 is the share of African American players drafted in any round of the draft. In Column 2, it is the share of African American players drafted in the first five rounds. In Column 3, it is the share of African American players drafted in later rounds (6+). The independent variable is the index of fan bias in attendance. * p<0.10, ** p<0.05, *** p<0.01. Standard errors in parentheses.