SUPPLEMENTAL ONLINE MATERIALS
Table of Contents

1. Additional details on perceived race coding

1.1. Coding process. We conceptualized players' race as a *perceived* race, not a self-identified race (Dixon and Maddox 2005, Maddox 2004, Wilkins et al. 2010), as our research question is about how others perceive the players and make decisions about them. To code each player's perceived race, we followed previous studies and made judgments mostly based on their appearance (e.g., Biegart et al. 2023, Clawson and Jett 2019, Clawson and Kegler 2000, MacLeod and Newall 2022, Zhang 2019).

Step 1. Each player was coded by two independent coders. The coders were instructed to first code based on their photos on the MLB website, using seven categories (1 = White non-Latinx; 2 = Black/ African American; 3 = Hispanic/ Latinx; 4 = Asian; 5 = Middle Eastern; 6 = Native American; 7 = Uncertain). The coders had the option to assign multiple categories for players that they perceived to be of multiracial descent. If available, they also examined supplementary information online, such as biographies, articles, and fan sites. However, it is worth noting that only a small portion of players had the relevant supplemental information online. This process ensured that 5,475 unique players in our sample were reviewed and coded by two independent coders.

Step 2. We then checked the agreement rate on the race between the two coders – how likely they were to code a player to be the same race. The agreement rate between the two coders is 91%. We also checked the agreement rate on racial minority status between the two coders – how likely they were to code a player to be a racial minority or not, as our main hypothesis regards minority players in general. The agreement rate was 95%.

Step 3. For the players that the two coders did not agree on (9%; N = 474), a third coder coded the race independently without knowing the previous coding. When the third coder agreed with one of the two coders, we followed that coding. The majority of coders agreed on 430 players (out of 474 players) and the overall agreement rate was 99%. When the third coder did not agree with either of the two coders (N = 44), the authors discussed and decided on the final coding. It is worth noting that these players consist of less than 1% of our sample. We also excluded three players (Billy Martin, Bobby Chouinard, John Maxwell) due to insufficient information to code their race.

The goal of this coding was to capture the *perception* of the players – whether a specific player was perceived to be a racial minority or not. Given the high agreement rate, we believe our coding method accurately captured the perception. We also conducted two robustness checks.

1.2. Robustness check with census data. Another common method to code race is to make inferences based on one's last name (Imai and Khanna 2016). If a person's last name is "Gonzalez", there is a high likelihood that the person is Hispanic/ Latinx. The United States Census Data has information that with each last name, what percentage of the people with the last name fall into each race category. Therefore, we first used the Python API ("ethnicolr") to get the census data probability for each last name of the players in our dataset. For example, the census data shows that in the United States, 95% of the people with the last name "Gonzalez" are racial and ethnic minorities, 94% are Hispanic, 0.4% are Black, and 1% are other races. We calculated this probability for each last name in our dataset – the probability for them to be a specific race and the probability for them to be racial and ethnic minorities.

We also calculated the race percentage of players who have a specific last name, based on our coding. That is, among the players with the last name "Gonzalez", what percent of them were coded as a racial minority by our coders? We then compared the two numbers – the census data probability and our coding percentage – by subtracting the former from the latter. For example, in our dataset, there were 27 players had the last name "Gonzalez." Among these players, 21 players were coded as Hispanic, 3 players were coded as Black, and 3 players were coded as White non-Latinx. Therefore, the difference for the last name "Gonzalez" as a racial and ethnic minority name is 6% (89% from our coding vs. 96% from census data).

On average, we found very small differences between the two numbers. Across all last names, there was only a 2% difference between the probability of a player with a specific last name being a racial and ethnic minority (census data) vs. the percentage of a player with a specific last name being coded as a racial minority (our coding). This indicates high consistency between our perceived race coding and census data last-name probability. To ensure that the results were not due to the specific API, we used an alternative API in R ("predictrace") and conducted the same analyses. The results were consistent.

1.3. Robustness check with available MLB statistics. As another robustness check, we compared our coded data with two publicly available datasets on MLB race statistics: Society for American Baseball Research (SABR) data and The Institute for Diversity and Ethics in Sports (TIDES) data. Both datasets do not contain data on individual players' race, but they have aggregate percentages of the player race for our time period (1988 – 2019): what percent of MLB players in 2019 were coded as White non-Latinx in their data? We thus collected the percentages from the two datasets and compared them with our coded data. We calculated the mean differences for each dataset: the % of racial and ethnic minority players reported by SABR (or TIDES) – the % of racial minority players resulted from our coding. The results indicate high agreement between the two datasets and our coded data. On average, we found 0.2% difference between SABR data and our coded data, and 2% difference between TIDES data and our coded data.

2. Additional details on status measurement

We constructed a composite status score using five different status indicators of each player: salary, awards, tenure, starting games, and celebrity status. Table S1 presents descriptive statistics and correlations for the five raw status indicators before scaling them within each team and year. Among the five indicators, we provide more details on the process of measuring and constructing three indicators – awards, tenure, and celebrity profiles.

2.1. Awards. Although there are a number of awards and honors in MLB, we chose the six awards listed as the "major awards" by Baseball Reference: Gold Glove (GG), Silver Slugger (SS), Most Valuable Player (MVP), Cy Young Award (CYA), Rookie of the Year (ROY), and All-Star nomination (AS). As mentioned in our manuscript, we assigned a weight to each award based on how valuable and difficult it is to win each award. First, the MVP award was assigned with the highest weight of four, as it was given to only two players per year to honor "the most important and useful player." Next, we assigned a weight of three to the following three awards because each of them was awarded to only one player in each league (and/or each position) to recognize performance in a specific domain. The CYA recognized a player who exhibited the best pitching performance, the GG award recognized a player who exhibited the best fielding performance, and the SS award was given to the best offensive player. We then assigned a weight of two to the AS nomination because, compared to the other awards, a significantly higher number of players (30 – 34 players per league) could receive the AS nomination. Lastly, we assigned the lowest weight of one to the ROY award, as it was awarded to the best rookie player in each league, who tended to have less experience and prestige compared to the players who won other awards. Indeed, experts tend to agree that the prestige associated with other awards, including the AS nomination, is generally considered to surpass that of the ROY award (Bleacher Report 2023).

Among all the 930 team-level observations, there were five observations in which no player had won any award (e.g., Atlanta Braves in 1990). For these five observations, there would be null values for the composite status scores during the computation due to the "division by zero" error (i.e., the maximum number of awards was zero for those observations). Therefore, we set the scaled score of awards to 0 for every player in those teams. The scaled zero scores distributed an equal award status for every player in the team and allowed us to calculate the within-status for players in these teams more accurately.

2.2. Robustness checks for awards. Once we collected the weighted values of all six awards, we calculated the *weighted* sum of the total number of awards won by each player. First, it is worth noting that we conducted a robustness check using the *unweighted* sum of the total number of awards won by each player as the award indicator. The results were listed in Table S2, which presented consistent results in direction and significance. Second, we also conducted a robustness check using the weighted sum with different weights given to the six awards. Following the comments from the review team, we weighed the Cy Young award, the MVP, and the Rockier of the Year award as the three equally higher-weighed awards, with a weight of two, than the All-star nomination, the Golden Glove award, and the Silver Slugger award, which are assigned with a weight of one. The results were listed in Table S3, which presented consistent results in direction and significance. These two additional analyses showed that the results in the manuscript were not driven by the specific weighing decision for the award variable.

2.3. Tenure. For each player's tenure, we defined it as the number of years since the player first appeared on the opening day roster in the major league as recorded by the USA Today Opening Day Roster. We used USA Today as the primary source of data as we believe the players on this opening day roster best align with full-time employees in a typical organization, unlike other players, such as those traded as free agents or called upon in mid-season. For every player, we obtained the debut year (T_0) when each player first appeared in the roster and calculated the tenure variable for the player in any later given year T as $(T - T_0 + 1)$. However, since the USA Today database only started in 1988, we needed to search the Baseball Reference 40-man roster for those players who first appeared in 1988 to update their debut year in case they appeared before 1988.

2.4. Robustness checks for tenure. We conducted several robustness checks. First, instead of calculating the continuous number of years since their debut, we counted the number of times they have appeared in the major league games using rosters from Baseball Reference. Our main effects remained consistent: The highest status of minority players predicted the number of minority players subsequently added to the team (b = -2.23, SE = 0.63, p < 0.001; Table S4). Next, due to the limitations that 1) the USA Today data started in 1988 and 2) Baseball Reference data did not have an opening day roster, our main analyses had to rely on both datasets. We conducted the analyses using alternative versions of the tenure variable by using either Baseball Reference or USA Today data only. When calculating the tenure variable using only the Baseball Reference rosters, the main effects remained consistent: The highest status of minority players predicted the number of minority players added to the team (b = -2.56, SE = 0.62, p < 0.001). Alternatively, we calculated the tenure variable using only the USA Today rosters by we set 1988 as their debut year for the players appeared in 1988. Again, our effects remain consistent: The highest status of minority players predicted the number of minority players subsequently added to the team (b = -2.66, SE = 0.61, p < 0.001).

Lastly, it is worth noting that tenure is often correlated with age. Indeed, on the individual level, the raw tenure variable and the age variable were highly correlated (r = 0.83, p < 0.001). Thus, we conducted a supplementary analysis, controlling for the age of both the highest-status minority player and the highest-status white player in each team each year in our main model. Again, the main effect remained consistent: The highest status of minority players still predicted the number of minority members added to the team (b = -2.57, SE = 0.62, p < 0.001).

2.5. Celebrity profile. For each player's celebrity profile, we followed Christie and Barling's work (2010) to collect the total number of articles on the magazine *Sports Illustrated* mentioning the player's name. Specifically, we first retrieved all the articles on *Sports Illustrated* by searching each player's name through the archival API (https://vault.si.com/search). We scraped the links and content of all the returned articles. Then for each player, we calculated the number of articles that mentioned the player's name in each year and used the result as the celebrity profile indicator.

3. Additional tests for FGLS regression

3.1. Likelihood ratio test. We used the likelihood ratio test to detect the presence of panel-level heteroskedasticity in all models. For each model, we conducted feasible generalized least squares (FGLS) regression with all the variables used in the model, including the dependent, independent and all the control variables. The regression was conducted twice without using the autocorrelation structure, and assuming the panel-level heteroskedasticity existed in the first model but did not exist in the second model. Then we obtained and stored these two models. As we conducted the FGLS regressions in the iterated way, the first model with heteroskedasticity produced maximum-likelihood parameter estimates. We thus conducted the likelihood ratio test to test the null hypothesis that there was the presence of homoskedasticity (instead of heteroskedasticity; i.e., the absence of heteroskedasticity) in our model by comparing these two models.

3.2. Wooldridge test. Similarly, we conducted the Wooldridge test to detect the presence of autocorrelation in each model. Analogous to the likelihood ratio test, we included all the variables in each model to conduct the Wooldridge test with a null hypothesis that there was no first-order autocorrelation in the longitudinal data used for the specific model.

Together, for each model reported in our paper and SOM, we conducted both the likelihood ratio test and the Wooldridge test to test the presence of heteroskedasticity and autocorrelation and decided whether we should include the heteroskedastic error structure or the auto-correlation structure in the regression. For each model, we included the panel-specific autocorrelation structure if the result of the Wooldridge test was significant. Similarly, we included the heteroscedasticity error structure if the result of the likelihood ratio test was significant. Here, we report the results of the tests for all FGLS regressions in our paper and SOM in Tables S5 and S6.

4. Additional analyses as robustness checks

4.1. FGLS regression analysis for changes in two years. As indicated in the manuscript, we considered the potential issue of having players whose contract length was longer than one year. To address this concern, instead of using the change in the number of racial minority players in the next one year, we used the changes in the next two years as our dependent variable. We conducted the FGLS regression with the same independent variable and control variables. The result showed a consistent pattern with our main analysis: The highest status of minority players predicted the number of minority players added to the team in the next two years (b = -4.01, SE = 0.83, p < 0.001; Table S7).

4.2. FGLS regression analyses for multiple high-status minority members. As indicated in the manuscript, the predictor variable was the *degree* of the highest status of minority players in our main analysis. We attempted to check the effect of having multiple high-status minority players. For each team, we calculated the *number* of minority players whose status was one-standard-deviation (1SD) above the mean status of the whole team. We then used the number of high-status minority players as our independent variable (M = 1.85, SD = 1.16) and conducted the regression analysis. The results provided additional support for our main hypothesis, suggesting that the teams with more high-status minority players hired fewer minority players than the teams with fewer high-status minority players (b = -0.61, SE = 0.09, p < 0.001; Table S8). The results were consistent in direction and significant when we used a different cut-off point (2SD above the mean status), suggesting that the teams with more high-status minority players hired fewer minority players (b = -0.55, SE = 0.15, p < 0.001).

4.3. FGLS regression analysis controlling for the change in the total number of players. As indicated in the manuscript, we considered a possibility that the change in the number of minority players is solely driven by the change in the total number of players. First, the change in the total number of players was indeed positively correlated with the change in the number of minority players (r = 0.46, p < 0.001), meaning that teams making more hires overall would naturally be more likely to hire more minority players. Thus, we conducted the same analysis, controlling for the change in the total number of players. The results remained consistent: The highest status of minority players predicted the number of minority players added to the team (b = -1.98, SE = 0.53, p < 0.001; Table S9).

4.4. FGLS regression analysis using a lagged model. As indicated in the manuscript, we used a change-score model for the main analysis. First, we chose a change-score model over a lagged model, because it better aligns with our theoretical focus on hiring outcomes for minority members from a longitudinal perspective (how many more or fewer minority members are added to a group over time) rather than focusing on the absolute level of diversity (how many minority members are present each year). Still, we conducted an additional analysis by using lagged model, where we used the number of minority players in the next year (Y_2) as the dependent variable and controlled for the number of minority players in the current year (Y_1). We conducted the regression analysis with the same control variables. The result remained consistent with our main analysis and suggested that the highest status of minority players predicted the number of minority players in the team in the following season (b = -1.70, SE = 0.61, p = 0.005; Table S10).

4.5. FGLS regression analysis using a change-score model controlling for the current number of minority members. For our main analysis, we used a change-score model. Based on the previous recommendations (Allison 1990, Sorjonen et al. 2021), we did not include the current number of minority members per team. This is because the current number of minority players was used to calculate the outcome variable: the change in the number minority players = the number of minority players in the next year (Y_2) – the number of minority players in the current year (Y_1). The same applies to the current percentage of minority players per team as it is technically the same variable as the current number of minority players due to the relatively fixed team size in MLB. Indeed, the percentage of minority players is significantly and highly correlated with the raw number of minority players (r = 0.93, p < 0.001), indicating that the two variables are almost statistically equivalent. We thus did not include it for the main analysis to avoid violating statistical principles. However, we did conduct an additional analysis controlling for the current number of minority players in the current year. The results remained consistent with our main analysis, suggesting that the highest status of minority players predicted the number of minority players added to the team (b = -1.70, SE = 0.61, p = 0.005; Table S11).

5. Additional details on player performance and position

5.1. Player performance. We measured players' performance using the Wins-Above-Replacement (WAR) statistics from Baseball Reference data (https://www.baseball-reference.com). Previous research suggested that the WAR is a reliable measure of individual players' performance (e.g., Baumer et al. 2015, Jensen 2013,Sanders et al. 2019). The definition of WAR of each player was derived from a replacement level win rate of a team per season – the win rate of a team with players who are all at replacement level. Specifically, the WAR of a player reflected the extra games that the team could win by having the player, compared to having another replacement player.

There are two types of WAR: pitcher WAR and batter WAR. The pitcher WAR was calculated through the runs allowed and saved, adjusted by the wins above average. The batter WAR was calculated through various runs including batting and base-running runs. As our analyses required a common performance measure for players across different positions, we followed the previous work (Baumer et al. 2015, Ehrlich et al. 2021) and constructed a total WAR by combining the pitcher WAR and the batter WAR. It is worth noting that the results were consistent when we used different versions of WAR (e.g., a primary WAR where we used pitcher WAR for pitchers and batter WAR for non-pitchers).

5.2. Player position. In the current dataset, we have 5,475 unique players who played in the 30 teams across 32 years. After excluding the 3 players whose races were not coded, these players yielded 26,053 observations where each observation annually recorded a player's information (e.g., status, position). First, using the whole sample, we present how the players were distributed between pitcher and non-pitcher positions (see Table S12). It is worth noting that racial minority players were underrepresented in pitcher position, compared to White players; there were 9,192 observations for White pitchers, while there were 3,260 observations for minority pitchers. To put it differently, among all observations for White players (N = 16,492), 56% were pitchers. In contrast, among all observations for minority players (N = 9,561), 34% were pitchers. Also, we present the position distribution for the highest-status players per team (see Table S13). We again found that the highest-status minority players were underrepresented in a pitcher position, compared to the highest-status White players. Among all observations for highest-status White players, 24% (N = 227) were pitchers. In contrast, among all observations for highest-status minority players, 6% (N = 58) were pitchers.

6. Experiment instructions and items

6.1. Annual statements. We pre-tested the two statements (N = 100) to ensure they conveyed a similar level of previous investment in diversity and inclusion ("How much effort do you think Company [A/ B] invested in diversity and inclusion?"; p = .234). Additionally, we ensured both statements were rated comparably in terms of readability and writing quality.

In the main experiment, we counterbalanced the presentation of the annual statements. That is, half of the participants read the first annual statement as the statement of Company A (high-status female executive) and the second one as the statement for Company B (low-status female executive). The other half of participants read the first annual statement as the statement of Company B and the second one was the statement of Company A. There was no significant difference between the two groups on our dependent variables.

Annual statement #1 (Company A/ B)
"*Dear Stakeholders,*
We are guided by a set of core values that define who we are and how we conduct business.
Innovation: We foster a culture of continuous improvement, providing our teams with the resources and freedom to explore new ideas and approaches. Our dedicated R&D division collaborates with external partners and invests in emerging technologies, driving innovation across our product and service offerings.
Customer Focus: Deeply understanding our customers' requirements, we have implemented robust feedback mechanisms to capture their insights and preferences. This enables us to develop tailored solutions that address their unique needs, provide exceptional experiences, and build lasting relationships based on trust and mutual success.
Diversity & Inclusion: We promote an inclusive environment where every individual is respected and has equal opportunities. Our Diversity & Inclusion initiatives encompass recruitment practices that ensure a diverse talent pool, inclusive leadership development programs, and employee resource groups that foster a sense of belonging and representation across our organization. Of all our employees, more than half identify as women, while 37% identify as people of color.
Sustainability: We embrace our responsibility to protect the environment and contribute to the well-being of society. Through our comprehensive sustainability strategy, we have implemented initiatives such as waste reduction programs, energy-efficient practices, responsible sourcing, and community engagement efforts that address social and environmental challenges.
Integrity: Upholding unwavering honesty, ethics, and transparency, we prioritize responsible governance and accountability in all our operations. Our robust compliance program ensures adherence to legal and regulatory requirements, safeguarding the trust and confidence of our stakeholders.
Together, we will continue to uphold these values, driving sustainable growth and making a positive impact on the world.
Sincerely,
President and Chief Executive Officer"

Annual statement #2 (Company B/ A)
"Dear Stakeholders,
We are delighted to present our core values.
Adaptability: We embrace change and proactively respond to evolving market dynamics. Our organization promotes a culture of learning and growth, encouraging our teams to explore new ideas, adapt to emerging technologies, and embrace innovation. We continuously invest in employee development programs to enhance adaptability and maintain our competitive edge.
Customer-Centricity: We prioritize understanding our customers' needs, expectations, and aspirations. Through customer feedback mechanisms, market research, and analytics, we develop tailored solutions and provide exceptional experiences. Our customer-centric approach ensures their satisfaction and

success.
Diversity, Equity, and Inclusion (DEI): We foster an inclusive environment that embraces individuals from diverse backgrounds, experiences, and perspectives. Our commitment to DEI is reflected in our recruitment practices, training programs, and leadership development initiatives that promote equal opportunities and ensure everyone feels valued and respected. Currently, 52% of all corporate roles are held by women, while one third of the roles are held by people of color.
Social Responsibility: We integrate sustainable practices into our operations, prioritize responsible sourcing, reduce our carbon footprint, and support local communities through philanthropic initiatives. Our commitment aligns our business objectives with the well-being of the planet and the communities we serve.
Accountability: We hold ourselves responsible for our actions, decisions, and their consequences. Our commitment is evident in our robust governance framework, adherence to legal and regulatory standards, and transparent reporting. We foster a culture of integrity and ethical conduct.
Together, we will embody these values, driving sustainable growth and delivering value to all.
Sincerely,
President and Chief Executive Officer"

6.2. Manipulations and items. We included the manipulation and exact items used in the study.
Instructions:
"*Please examine the profiles of the newly appointed executives of the two companies. All of them were hired this year. All of them are external hires from other companies. Each profile has information about the executive's position (the current position in the company), salary (the annual salary from the company), industry experience (the number of years the executive has worked in the industry), awards (the awards the executive has received recently), and celebrity status (the number of followers the executive has on LinkedIn)*."

Company A.

Name	Gender	Title	Salary	Tenure	Award	Followers
Mrs. Johnson	Female	CFO	$2,430,000	15 years	Executive Leadership Award, Business Brilliance Award	2,053,206
Mr. Miller	Male	CIO	$1,260,000	13 years	Excellent Entrepreneur of the Year	1,001,602
Mr. Jones	Male	CTO	$930,000	11 years	None	6,538
Mr. Davis	Male	COO	$300,000	7 years	None	2,273

Company B.

Name	Gender	Title	Salary	Tenure	Award	Followers
Mrs. Anderson	Female	CIO	$280,000	8 years	None	3,011
Mr. Smith	Male	CTO	$2,016,000	17 years	Innovative Leadership Award, Digital Leadership Award	2,011,745
Mr. Brown	Male	CFO	$1,850,000	12 years	Leader Impact Award	1,321,689
Mr. Williams	Male	COO	$890,000	10 years	None	7,490

Perceived status:

"Rate each executive on the extent to which you think the executive has high or low status in Company [A/ B]. Status means "respect, admiration, and prominence that an individual enjoys in the eyes of others." Please rate each individual in Company [A/ B] from lowest (1) to highest (7) status."

Previous effort:
"How much resource do you think Company [A/ B] invested in hiring a female executive?"
"How much effort do you think Company [A/ B] invested in hiring a female executive?"

Future effort for hiring:
"How much more or less resources do you think Company [A/ B] needs to invest in hiring a female executive in the future?"
"How much more or less effort do you think Company [A/ B] needs to invest in hiring a female executive in the future?"
"How much more or less money do you think Company [A/ B] needs to invest in hiring a female executive in the future?"

Other variables:
Gender, age, race, education

Table S1. Descriptive statistics and correlations for five status indicators

	Mean	S.D.	1	2	3	4	5
1. Salary	2458635.68	3917921.27	-				
2. Starting games	48.97	51.26	0.24	-			
3. Tenure	5.66	4.31	0.41	0.11	-		
4. Awards	0.31	1.11	0.24	0.34	0.08	-	
5. Celebrity status	2.83	4.81	0.34	0.32	0.29	0.44	-

Note. All variables are in raw values. The salary is in USD. All correlations are significant at $p < 0.001$

Table S2. Regression analysis estimating the change in the number of minority players, using unweighted awards

	b	*S.E.*
Minority highest status	-2.64*** [-4.36]	0.61
Minority average status	3.24** [2.88]	1.12
White highest status	-0.44 [-0.72]	0.61
White average status	-0.01 [-0.01]	1.29
Minority average performance	0.84*** [4.89]	0.17
White average performance	-0.01 [-0.05]	0.24
Team winning percentage	-4.47* [-2.47]	1.81
Performance of the highest-status minority player	-0.07 [-1.87]	0.04
Performance of the highest-status White player	0.03 [0.75]	0.03
Constant	2.84** [3.03]	0.94

Note. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. FGLS regression with panel-specific AR (1) autocorrelation structure. R^2 statistics are not available in FGLS (Greene 2000). Z-statistics are in brackets. The unweighted sum of the awards was used as one of the status indicators.

Table S3. Regression analysis estimating the change in the number of minority players, using different award weights

	b	*S.E.*
Minority highest status	-2.64*** [-4.36]	0.61
Minority average status	3.26** [2.91]	1.12
White highest status	-0.42 [-0.70]	0.61
White average status	0.07 [0.05]	1.29
Minority average performance	0.83*** [4.84]	0.17
White average performance	-0.02 [-0.07]	0.24
Team winning percentage	-4.47* [-2.47]	1.81
Performance of the highest-status minority player	-0.07 [-1.76]	0.04
Performance of the highest-status White player	0.02 [0.71]	0.03
Constant	2.81** [2.99]	0.94

Note. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. FGLS regression with panel-specific AR (1) autocorrelation structure. R^2 statistics are not available in FGLS (Greene 2000). Z-statistics are in brackets.

Table S4. Regression analysis estimating the change in the number of minority players, calculating tenure variable as the number of appearances

	b	*S.E.*
Minority highest status	-2.23*** [-3.53]	0.63
Minority average status	2.70* [2.27]	1.19
White highest status	-0.21 [-0.33]	0.64
White average status	-0.83 [-0.61]	1.36
Minority average performance	0.92*** [5.27]	0.17
White average performance	0.02 [0.08]	0.24
Team winning percentage	-4.71* [-2.53]	1.86
Performance of the highest-status minority player	-0.08* [-2.10]	0.04
Performance of the highest-status White player	0.03 [0.87]	0.03
Constant	2.76** [2.83]	0.98

Note. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. FGLS regression with panel-specific AR (1) autocorrelation structure. R^2 statistics are not available in FGLS (Greene 2000). Z-statistics are in brackets.

Table S5. Likelihood ratio tests for detecting the presence of heteroskedasticity

Model	LR chi2(29)	*p*	Heteroscedasticity
Table 3 Model (1)	27.75	0.532	No
Table 3 Model (2)	37.46	0.135	No
Table 3 Model (3)	38.63	0.109	No
Table 4 Model (1)	37.82	0.127	No
Table 4 Model (2)	36.17	0.169	No
Table 5 Model (1)	46.34	0.022	Yes
Table 5 Model (2)	45.59	0.026	Yes
Table S2	37.59	0.132	No
Table S3	37.05	0.145	No
Table S4	39.26	0.097	No
Table S7	27.38	0.551	No
Table S8	35.82	0.179	No
Table S9	44.34	0.034	Yes
Table S10	25.75	0.639	No
Table S11	25.75	0.639	No

Note. H0: no heteroscedasticity. Tables 3, 4, 5 are included in the manuscript.

Table S6. Wooldridge test for detecting the presence of autocorrelation

Model	F(1, 29)	*p*	Autocorrelation
Table 3 Model (1)	50.353	0.000	Yes
Table 3 Model (2)	9.905	0.004	Yes
Table 3 Model (3)	8.276	0.008	Yes
Table 4 Model (1)	29.666	0.000	Yes
Table 4 Model (2)	2.699	0.111	No
Table 5 Model (1)	8.071	0.008	Yes
Table 5 Model (2)	0.389	0.538	No
Table S2	10.564	0.003	Yes
Table S3	10.729	0.003	Yes
Table S4	7.871	0.009	Yes
Table S7	218.009	0.000	Yes
Table S8	17.405	0.000	Yes
Table S9	5.562	0.025	Yes
Table S10	178.208	0.000	Yes
Table S11	176.403	0.000	Yes

Note. H0: no first-order autocorrelation. Tables 3, 4, 5 are included in the manuscript.

Table S7. Regression analysis estimating the change in the number of minority players within two years

	B	*S.E.*
Minority highest status	-4.01*** [-4.86]	0.83
Minority average status	5.79*** [3.83]	1.51
White highest status	1.36 [1.62]	0.84
White average status	-2.96 [-1.59]	1.86
Minority average performance	0.96*** [4.08]	0.23
White average performance	0.17 [0.52]	0.33
Team winning percentage	-5.59* [-2.24]	2.50
Performance of the highest-status minority player	0.00 [0.01]	0.05
Performance of the highest-status White player	-0.03 [-0.57]	0.05
Constant	2.72* [2.06]	1.32

Note. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. FGLS regression with panel-specific AR (1) autocorrelation structure. R^2 statistics are not available in FGLS (Greene 2000). Z-statistics are in brackets.

Table S8. Regression analysis estimating the change in the number of minority players as a function of the number of high-status minority players

	b	*S.E.*
Number of high-status minority players	-0.61*** [-6.86]	0.09
Number of high-status White players	0.08 [0.96]	0.08
Minority average status	5.52*** [4.92]	1.12
White average status	-2.71* [-2.03]	1.33
Minority average performance	0.65*** [4.21]	0.16
White average performance	-0.06 [-0.26]	0.23
Team winning percentage	-3.71* [-2.05]	1.81
Constant	0.94 [1.23]	0.76

Note. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. FGLS regression with panel-specific AR (1) autocorrelation structure. R^2 statistics are not available in FGLS (Greene 2000). Z-statistics are in brackets. We did not include the highest status performances for minority and White players as this model involves multiple high-status players. The results for the main effect were consistent in direction and significant when we controlled for the two variables (b = -0.61, SE = 0.09, $p < 0.001$).

Table S9. Regression analysis estimating the change in the number of minority players, controlling for the change in the total number of players

	b	*S.E.*
Minority highest status	-1.98*** [-3.73]	0.53
Minority average status	1.14 [1.15]	1.00
White highest status	0.18 [0.34]	0.53
White average status	-1.46 [-1.29]	1.13
Minority average performance	0.70*** [4.62]	0.15
White average performance	-0.39 [-1.87]	0.21
Team winning percentage	-1.46 [-0.92]	1.59
Performance of the highest-status minority player	-0.08* [-2.19]	0.03
Performance of the highest-status White player	0.02 [0.52]	0.03
Change in the total number of players	0.32*** [15.27]	0.02
Constant	2.10* [2.58]	0.81

Note. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. FGLS regression with panel-specific AR (1) autocorrelation structure and correcting for panel heteroskedasticity. R^2 statistics are not available in FGLS (Greene 2000). Z-statistics are in brackets.

Table S10. Regression analysis estimating the number of minority players per team in the following year using a lagged model

	b	*S.E.*
Minority highest status	-1.70** [-2.80]	0.61
Minority average status	2.41* [2.17]	1.11
White highest status	-1.89** [-3.14]	0.60
White average status	1.21 [0.94]	1.29
Minority average performance	0.70*** [4.19]	0.17
White average performance	-0.44 [-1.87]	0.24
Team winning percentage	-1.03 [-0.57]	1.80
Performance of the highest-status minority player	-0.07 [-1.82]	0.04
Performance of the highest-status White player	0.03 [0.91]	0.03
Current number of minority players	0.57*** [19.97]	0.03
Constant	6.51*** [6.76]	0.96

Note. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. FGLS regression with panel-specific AR (1) autocorrelation structure. R^2 statistics are not available in FGLS (Greene 2000). Z-statistics are in brackets.

Table S11. Regression analysis estimating the change in the number of minority players, controlling for the current number of minority players

	b	*S.E.*
Minority highest status	-1.70** [-2.80]	0.61
Minority average status	2.41* [2.17]	1.11
White highest status	-1.89** [-3.14]	0.60
White average status	1.21 [0.94]	1.29
Minority average performance	0.70*** [4.19]	0.17
White average performance	-0.44 [-1.87]	0.24
Team winning percentage	-1.03* [-0.57]	1.80
Performance of the highest-status minority player	-0.07 [-1.82]	0.04
Performance of the highest-status White player	0.03 [0.91]	0.03
Current number of minority players	-0.43*** [-14.92]	0.03
Constant	6.51*** [6.76]	0.96

Note. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. FGLS regression with panel-specific AR (1) autocorrelation structure. R^2 statistics are not available in FGLS (Greene 2000). Z-statistics are in brackets.

Table S12. Status information of pitchers and non-pitchers of the whole sample

	White pitcher (N = 9,192)	White non-pitcher (N = 7,300)	Minority pitcher (N = 3,260)	Minority non-pitcher (N = 6,301)
Salary	2,194,152	2,412,576	2,511,669	2,871,811
Tenure	5.40	5.76	5.19	6.16
Games started	10.93	79.60	10.01	86.10
Awards score	0.17	0.37	0.15	0.51
Media coverage	2.38	3.00	2.08	3.68
Performance	0.97	1.23	0.96	1.38

Note. All values are average values.

Table S13. Status information of pitchers and non-pitchers of the highest-status player sample

	White pitcher (N = 227)	White non-pitcher (N = 703)	Minority pitcher (N = 58)	Minority non-pitcher (N = 872)
Salary	9,463,513	7,078,691	9,328,579	7,351,363
Tenure	11.25	8.52	10.12	9.11
Games started	27.11	131.29	20.88	129.09
Awards score	2.02	2.27	1.22	2.29
Media coverage	12.98	9.11	5.76	10.25
Performance	3.92	3.53	2.43	3.18

Note. All values are average values. The highest-status player sample consists of only the players with the highest status score within their team.