Appalachian
STATE UNIVERSITY

Department of Economics Working Paper

Number 22-08 | September 2022

Positional Segregation and Career Length in Major League Baseball in the 1990s

Peter Groothuis
Appalachian State University

Johnny Ducking
North Carolina A&T State University

Department of Economics
Appalachian State University
Boone, NC 28608
Phone: (828) 262-2148
Fax: (828) 262-6105
www.business.appstate.edu/economics

Positional Segregation and Career Length in Major League Baseball in the 1990s

Johnny Ducking
Associate Professor of Economics
Department of Economics
North Carolina A&T State University
Greensboro, NC 27411
jcduckin@ncat.edu

Peter A. Groothuis
Professor of Economics
Department of Economics
Appalachian State University
Boone, NC 28607
groothuispa@appstate.edu

Abstract: The racial makeup of Major League Baseball (MLB) has changed over time with the percentage of Blacks in the league peaking at eighteen percent in the late 1980s and then declining to only seven percent today. While the percentage of Hispanics has grown steadily over time reaching fifteen percent in 1990 and continuing to grow to twenty-eight percent today. To provide insight into both the decline in Black players and rise in Hispanic players, we use panel data from 1990 to 2004 to explore both the racial and ethnic positional segregation and career length of players. When it comes to career length, we find that foreign born Hispanic players have a higher probability of exit while Hispanic players who were born in the United States or the territory of Puerto Rico have no difference in the probability of exit than non-Hispanic White players. In addition, we find that non-Hispanic Black players have no difference in the likelihood of exit than non-Hispanic White players when there are no controls for the position played. However, we do find that Black non-Hispanic players have a lower probability of exiting the league when position controls are included. Focusing on position, we find that non-Hispanic Black players are most likely to be in the outfield, least likely to be a catcher, and outfielders have shorter careers than catchers. Lastly, we find that Hispanic players are most likely to be a shortstop and least likely to be a catcher.

Keywords: Discrimination, Training, Sports, Major League Baseball (MLB)

Parents are doling out up to $5,000 to have their sons play on travel teams with multiple sets of fancy uniforms, up to $500 to attend showcase camps in which they walk away with promotional CD-ROMs of their son and up to $60 an hour once or twice a week in the off-season to have Johnny take private lessons. The young Black without access to that kind of intensive training is hopelessly behind the learning curve of a game that is difficult to grasp. "We've lost them by age 13," says DeJon Watson, director of professional scouting for the Indians.

Tom Verducci, Sports Illustrated

I. Introduction

Since the desegregation of baseball in 1947 the percentage of players who are Black and Hispanic have followed different paths. The percentage of Black players first increased reaching the peak in the mid to late 1980s at eighteen percent of players then declined to just seven percent today. The percentage of Hispanic players, however, have continued to increase over time from less than one percent in 1947 to over twenty-nine percent today. To gain insights on both the fall in Black players and continued rise of Hispanic players, we analyze a panel of baseball players from 1990 through 2004. We focus both on positional assignment and career length to discern if there is either positional discrimination or exit discrimination in major league baseball.

Focusing on career length, we find that foreign born Hispanic players have a higher probability of exit while non-Hispanic Black players have no difference in the probability of exit than non-Hispanic White players when there are no controls for position included in the analysis. We do, however, find that non-Hispanic Black players have a lower probability of exiting the league when position controls are included while Hispanic players still have a higher probability of exiting the league after controlling for position.

When focusing on position, we find that there is clear evidence of positional assignment by racial characteristic after controlling for height, weight, and offensive performance statistics. We find that non-Hispanic Black players are most likely to be in the outfield and least likely to

be a catcher. Hispanic players are most likely to be a shortstop and least likely to be a catcher. Non-Hispanic White players are most likely to be catchers and first basemen.

II. Literature

Becker (1971) identifies three types of discrimination; employer, coworker, and customer, and points out that individuals who are in the group that is not preferred by consumers may earn less than individuals in the preferred group. Several studies have examined discrimination in sports over the past fifty years. Kahn (2000) indicates that productivity and compensation data along with the ability to examine the impact of new rules makes professional sports a great place to test hypotheses about labor markets. Scully (1973), Kahn (1991 and 2000), and Simmons (2021) discuss studies of discrimination in professional sports and all identify some evidence of discrimination in professional sports.

Kahn and Sherer (1988), Gius and Johnson (1998), Hill (2004), Kahn and Shah (2005), and Johnson and Minuci (2020) use data from the National Basketball Association (NBA) to test if there is evidence of discrimination in professional sports. Kahn and Sherer (1988) find that Blacks earn about 20% less than whites and that replacing a Black player with a white player increases home attendance. Gius and Johnson (1998) and Hill (2004) find no evidence of discrimination against Black players in the NBA while Kahn and Shah (2005) find evidence of discrimination against marginal nonwhite players. Johnson and Minuci (2020) uses NBA data on free agents and find evidence that Blacks are paid substantially less than their counterparts. Hoang and Rascher (1999) and Groothuis and Hill (2013) look at exit discrimination in the NBA. Hoang and Rascher (1999) finds that the expected career length of a Black player is about two years shorter than the expected career length for a White player which leads to Black players earning roughly $800,000 less than White players during their careers. Groothuis and Hill (2013) find no

evidence of exit discrimination but they find that White players earn sixteen to twenty percent more over their careers. Groothuis and Hill (2004) use panel data from 1989 to 1999 and find no evidence of exit discrimination in the NBA.

Mogull (1973), Kahn (1992), Keefer (2013), Burnett and Van Scyoc (2013 and 2015), and Ducking et al. (2014 and 2017) look at professional football to see if there is evidence of salary or compensation discrimination. Mogull (1973) finds no evidence of racial discrimination. Kahn (1992) finds a small difference in the salaries of Whites and nonwhites with Whites earning slightly more. Keefer (2013), Burnett and Van Scyoc (2013 and 2015), and Ducking et al. (2017) all utilize quantile regression analysis and the Oaxaca-Blinder decomposition developed by Oaxaca (1973) and Blinder (1973) to see if discrimination exists in professional football. Burnett and Van Scyoc (2013 and 2015) restricts their data to rookies and find no evidence of discrimination. Keefer (2013) and Ducking et al. (2017) both find evidence of the discrimination against Black linebackers. Ducking et al. (2014) looks at career earnings and finds no evidence of racial discrimination in the NFL and Ducking et al. (2015) looks at six positional groups and find no evidence of exit discrimination in the NFL.

Mogull (1975) finds that Black players are not discriminated against and receive higher salaries on average due to their superior performance. Palmer and King (2006) explore salary discrimination in Major League Baseball by looking at high, medium, and low salary groups. They find evidence of discrimination against Blacks and Hispanics in the low salary group. Holmes (2011) uses quantile regression analysis to look at salary discrimination in Major League Baseball by restricting the sample to free agents. He uses nine years of data and finds evidence of salary discrimination against Blacks in the lower half of the salary distribution.

Jibou (1988) and Groothuis and Hill (2008) both look at exit discrimination in major league baseball. Jibou finds that Blacks experienced shorter careers but Hispanics did not experience shorter careers. Groothuis and Hill (2008) use panel data from 1990 to 2004 and find no evidence of exit discrimination in Major League Baseball.

Medoff (1977) find that there are higher skill and development costs for central positions of pitcher, catcher, shortstop, second base, and third base. He indicates that the order of positions by decreasing relative cost is pitcher, catcher, shortstop, second base, and third base. Medoff (1986) attributes positional segregation to the inferior socioeconomic status of Blacks, differential skill, and development cost. Smith and Leonard (1997) described stacking as the placement of Black baseball players in non-central positions more often than they are placed in central positions. One explanation that they offered for stacking is that the individuals in charge do not believe that Black players have the ability to perform at thinking positions. Another explanation that they offered is that Blacks may want to play in the outfield because the players that they admire play in the outfield. Margolis and Piliavin (1999) indicate that part of the reason why Blacks are in the outfield is due to a job-related skill such as speed. Sack et al. (2005) uses multinomial logistic regression analysis and find evidence of occupational segregation in Major League baseball, which is referred to as stacking.

III. Data

We utilize the same data on hitters as Groothuis and Hill (2008) because this dataset allows us to take a historical look at Major League Baseball prior to the decline of the percentage of Blacks in the league to roughly seven percent. Our data is a 15-year panel that includes information

on all positions in Major League Baseball except pitchers from 1990 to 2004.[1] Both stock and flow samples are included in our analysis. The stock sample is left-censored and consists of careers that started prior to 1990. These data are easily included because we know how many years each player played in the Majors prior to 1990. The flow sample includes all careers that start between 1990 and 2004. This sample captures many short careers in Major League Baseball. Including only flow data however would allow for no careers longer than 15 years, which is the length of our panel. As with most panels, our data is also right-censored where many careers were ongoing when our sample ended in 2004. Our data includes both stock and flow observations that are right-censored. We use the technique developed by Berger and Black (1998) to estimate a duration model of stock and flow data.

Our data include information on player characteristics and performance. The player characteristics included in the data are age, experience, height, weight, dummy variables that capture race and ethnicity, and dummy variables for the player's position. The Hispanic variable in our data characterizes players of Hispanic descent. The White variable in our data characterizes non-Hispanic White players and the use of the term White in the remainder of this paper refers to non-Hispanic Whites. The Black variable in our data characterizes non-Hispanic Black players and the use of the term Black in the remainder of this paper refers to non-Hispanic Blacks. The Hispanic players in our data are from the USA, Puerto Rico, Dominican Republic, Colombia, Cuba, Curacao, Mexico, Nicaragua, Panama, and Venezuela.[2] Our performance data include games played, slugging average, stolen base percentage, runs batted in per game, walks

[1] We exclude hitter observations when the player appears in less than 20 games in the season. Major League Baseball roster limits are expanded for 25 to 40 from September 1st until the end of the season. In addition, we exclude all Asian born players because in the 1990s these made up just a small portion of players in the league.

[2] Players from Colombia, Cuba, Curacao, Mexico, Nicaragua, Panama, and Venezuela are grouped together to create the Hispanic other category due to small sample sizes.

per game, strikeouts per game, hit by pitch per game, sacrifice flies per game, caught stealing per game, and runs scored per game.

Table 1A: Means of Variables

Variables	White	Black	Hispanics
Age	29.55	29.82	28.12
Experience	4.61	5.81	4.31
Slugging Average	0.394	0.404	0.397
Runs Batted In per game	0.369	0.386	0.386
Walks per game	0.295	0.313	0.266
Strike-Outs per game	0.540	0.552	0.536
Stolen Bases per game	0.038	0.095	0.056
Games	93.03	102.54	98.46
Height	72.92	72.47	72.09
Weight	193.77	193.38	183.53
Number observations	3795	1621	1728

Table 1B: Means of Variables for Hispanic by Place of Origin

Variables	Hispanic USA	Hispanic Puerto Rico	Hispanic Dominican Republic	Hispanic Other
Age	29.01	28.57	27.68	27.79
Experience	4.28	4.95	4.31	3.74
Slugging Average	0.407	0.407	0.387	0.395
Runs Batted In per game	0.397	0.410	0.368	0.383
Walks per game	0.291	0.280	0.256	0.251
Strike-Outs per game	0.467	0.550	0.559	0.535
Stolen Bases per game	0.051	0.044	0.066	0.056
Games	98.16	99.88	97.36	98
Height	72.53	71.96	71.80	72.34
Weight	186.04	189.11	177.71	184.74
Number observations	266	408	599	455

We report the means by race and ethnicity in table 1A. In comparison to Whites and Hispanics, Blacks in our data are older, have more experience, a higher slugging average, more runs batted in per game, more walks per game, more strikeouts per game, more stolen bases per game, and more games played. In comparison to Blacks and Hispanics, Whites in our data are both taller and weigh more. We report the means for Hispanic by nation of origin in table 1B. Hispanics from Puerto Rico have more experience, runs batted in per game, and play more

games. Hispanics from the USA have more walks per game. Hispanics from the Dominican Republic have higher strikeouts per game and more stolen bases per game per game.

We report the proportion of players at each position by race and ethnicity in table 2A. Whites have a higher percentage of catchers, first basemen, second basemen, third basemen and utility fielders. In our data, 77.1% of the catchers are White, 64.0% of the first basemen are White, 48.7% of the second basemen are White, 75.3% of the third basemen are White, and 61.8% of the utility fielders are White. Hispanics have a higher percentage of short stops with 53.4% and Blacks have a higher percentage of outfielders with 43.1%. We report the proportion of Hispanics at each position by nation of origin in table 2B. Hispanics from Puerto Rico have a larger proportion of catchers. Hispanics from the Dominican Republic have a larger proportion of second basemen, third basemen, short stops, outfielders, and utility fielders. Other Hispanics have a larger proportion of first basemen.

Table 2A: Proportion of Position by Race

Position	White	Black	Hispanic
Catcher	.771	.024	.205
First Base	.640	.197	.162
Second Base	.487	.215	.297
Third Base	.753	.086	.162
Short Stop	.350	.113	.534
Outfield	.365	.431	.203
Utility Fielder	.618	.109	.273

Table 2B: Proportion of Hispanic Position by Nation of Origin

Position	Hispanic All	Hispanic USA	Hispanic Puerto Rico	Hispanic Dominican Republic	Hispanic Other
Catcher	.205	.016	.093	.035	.060
First Base	.162	.033	.004	.048	.076
Second Base	.297	.028	.100	.121	.049
Third Base	.162	.046	.023	.056	.038
Short Stop	.536	.067	.093	.221	.155
Outfield	.203	.038	.049	.064	.051
Utility Fielder	.273	.041	.046	.135	.052

Nonparametric Estimates of Career Duration

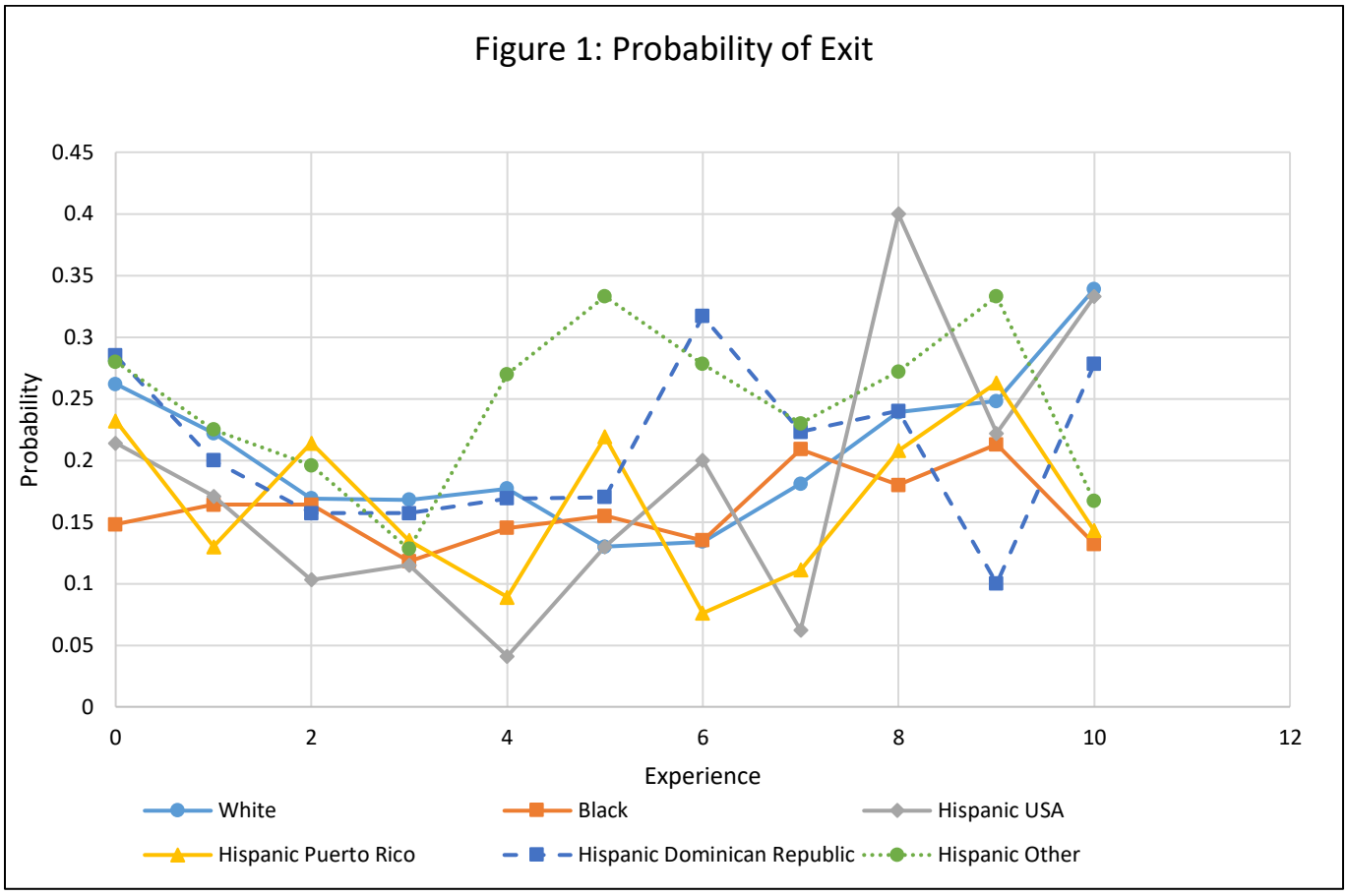
To help understand career duration in Major League Baseball, we calculate yearly hazard functions as $h_t = d_t/n_t$ where d, is the number of players who end their career in year t and n is the number of players at risk of ending their career in year t. The hazard rate can be interpreted as the probability players exit Major League Baseball given that they have survived up to some level of tenure. In table 3 and figure 1 we show the yearly hazard rates for Whites, Blacks, and Hispanics from different places of origin. Blacks have the lowest probability of exit when they enter Major League Baseball with no experience, and they also have the lowest probability of exit when they have ten years of experience. This is consistent with the hypothesis that only the best Black players are given a chance to play Major League Baseball, therefore once they make it to Major League Baseball they are less likely to exit.

Table 3: Probability of Exit

Experience	White	Black	Hispanic USA	Hispanic Puerto Rico	Hispanic Dominican Republic	Hispanic Other
0	0.262	0.148	0.214	0.232	0.285	0.28
1	0.222	0.164	0.171	0.13	0.2	0.225
2	0.169	0.164	0.103	0.214	0.157	0.196
3	0.168	0.118	0.115	0.135	0.157	0.128
4	0.177	0.145	0.041	0.089	0.169	0.27
5	0.13	0.155	0.13	0.219	0.17	0.333
6	0.134	0.135	0.2	0.076	0.317	0.278
7	0.181	0.209	0.062	0.111	0.223	0.23
8	0.239	0.18	0.4	0.208	0.24	0.272
9	0.248	0.213	0.222	0.263	0.1	0.333
10	0.339	0.132	0.333	0.143	0.278	0.167

Foreign born Hispanics in the Dominican Republic or elsewhere as indicated by the other category which includes Colombia, Cuba, Curacao, Mexico, Nicaragua, Panama, and Venezuela all have a higher hazard rate particularly early in their careers suggesting that foreign born

Hispanics are the least likely to stay in the major league once promoted. USA Hispanic and Puerto Rico born players, however, do not have higher hazard rates than White players early in their careers.



Semi-parametric Estimates of Career Duration

We estimate semi-parametric hazard functions following Berger and Black (1998), Groothuis and Hill (2004 and 2008), and Berger, Black, and Scott (2004). We calculate our hazard model as a discrete random variable because we have season level data. We follow the notation of Berger, Black, and Scott (2004) to understand how stock data influences a likelihood function. Suppose the probability mass function (pmf) of durations is defined as $f(t,x,\beta)$, where t is the duration of the career, x is a vector of performance and personal characteristics, and β is a

vector of parameters. Now denote F(t,x,β) as the cumulative distribution function; then the probability that a career lasts at least t° years is simply 1 - F(t°,x,β). If we define the hazard function as h(t,x,β) ≡ f(t,x,β) / S(t,x,β) where S is the survivor function,

$S(t, x, \beta)=\prod_{i=1}^{t-1}[1-h(i, x, \beta)]$ and apply the definition of conditional probabilities, we may express the pmf as

$$f(t_i, x_i, \beta) = \prod_{j=0}^{t_i-1}[1 - h(j, x_i, \beta)]h(t_i, x_i, \beta). \qquad (2)$$

If we have a sample of n observations, $\{t_1, t_2, \ldots, t_n\}$, the likelihood function of the sample is

$$L(\beta) = \prod_{i=1}^{n} f\ (t_i, x_i, \beta) = \prod_{i=1}^{n}\left(\prod_{j=1}^{t_i-1}[1 - h(j, x_i, \beta)]\, h(t_i, x_i, \beta)\right). \qquad (3)$$

Let the set A be the set of all observations where the players' careers are completed and the set B be the set of all observations where the careers are right censored. For the set of right-censored observations, all we know is that the actual length of the career is greater than t_i, the observed length of the career up through the last year. Because we know that the actual length of the career is longer than we observe then the contribution of these observations to the likelihood function is just the survivor function (S).

Let the set C be the set of careers that were in progress when data collection began. For these observations, we know that the career i has lasted for r years before the panel begins so the likelihood must be adjusted by the conditional probability of the career having length r. Some stock-sampled observations may be right-hand censored. Let the set D be the set of all stock-sampled observations that are also right-hand censored. An example of a career that is both right and left censored would be a player that starts his career prior to 1990 and ends his career after 2004. Taking into account all four sets: A, B, C, and D the likelihood function becomes

$$L(\beta)=\prod_{i\in A}\left(\prod_{j=1}^{t_i-1}[1-h(j,x_i,\beta)]h(t_i,x_i,\beta)\right)\times\prod_{i\in B}\left(\prod_{j=1}^{t_i-1}[1-h(j,x_i,\beta)]\right)$$
$$\times\prod_{i\in C}\left(\prod_{j=r_i}^{t_i-1}[1-h(j,x_i,\beta)]\right)h(t_i,x_i,\beta)\times\prod_{i\in D}\left(\prod_{j=r_i}^{t_i-1}[1-h(j,x_i,\beta)]\right) \quad (4)$$

In equation (4) the contribution of censored, stock-sampled observations to the likelihood function is strictly from the last two terms; such observations simply provide information about the survivor function between (r,t).

We, as Berger, Black and Scott (2004), have expressed the likelihood function as a function of the hazard functions. All that remains is to specify the form of a hazard function and estimate by means of maximum likelihood estimation. As the hazard function is the conditional probability of exiting MLB given that the MLB career lasted until the previous season, the hazard function must have a range from zero to one. In principle, any mapping with a range from zero to one will work. Cox (1972) recommends

$$\frac{h(t,x,\beta)}{1-h(t,x,\beta)}=\frac{h_t}{1-h_t}e^{x\beta}=\exp(\gamma_t+x\beta), \quad (5)$$

which is simply the logit model with intercepts that differ by time periods. The term h_t is a baseline hazard function, which is common to all. The $x\beta$ term, determined by the player's personal and productivity characteristics, shifts the baseline hazard function, but it affects the baseline hazard function in exactly the same way each period. Berger and Black (1998) consider other hazard functions and find that the results are relatively robust across various specifications of the hazard function.

Because the players in the panel have varying degrees of job tenure prior to the beginning of the panel, we identify the hazard function for both long and short careers. The disadvantage

to this approach is that the vector γ_t of equation (5) can be very large. In our study it would require 25 dummy variables. We also run into problems with the Cox technique because we have too few players who have long careers. To simplify the computation of the likelihood function and be able to keep the long careers, we simply approximate the γ_t vector with a 5th order polynomial of the players' tenure in MLB, which reduces the number of parameters to be estimated from 25 to 5. Thus, the hazard function becomes

$$\frac{h(t,x,\beta)}{1-h(t,x,\beta)} = \Phi(t)\, e^{x\beta} = \exp(\phi(t) + x\beta)\,, \qquad (6)$$

where $\phi(t)$ is a 5th order polynomial of the player's tenure in MLB. The 5th order polynomial therefore includes tenure to the first, second, third, fourth and fifth powers. We choose the Taylor series approximation technique over using tenure dummies due to the small number of observations for high tenures. This method provides a very flexible specification of the baseline hazard, but does impose more restrictions than Cox's model.[3]

In table 4 we report the results from the semiparametric analysis of career duration. In model (1) there are no controls for position while in model (2) there are controls for positions. The omitted race category is White and the omitted position in model (2) is the catcher position. In both specifications we find that the sign of the coefficients on the performance variables of slugging percentage, walks per game, stolen bases per game and the number of games played are all negative and significant showing that in both models that increases in these performance measures decrease the likelihood of exit. The sign of the coefficients on strikeouts per game are positive and significant in both models indicating that this negative performance measure

[3] The results do not change when higher order polynomials of the sixth and seventh power are included indicating that a fifth order polynomial is flexible enough to capture the influence of the baseline hazard.

increases the likelihood of exit. We also find that the higher the age of a player, the higher the likelihood of exit.

When focusing on ethnicity, we find that the signs of the coefficients on Hispanics from the Dominican Republic and Hispanics in our other category are both positive and significant at the 95% level in both models indicating that these foreign-born Hispanics are more likely to exit in comparison to non-Hispanic White players. To convert the coefficients into a percentage and focus on the magnitude of the effect we use 100[exp(P) - 1]. This conversion gives us the percentage difference in hazard rates between Whites and foreign-born Hispanics. Using this method, we find that Hispanic players who were born in the Dominican Republic are 46% more likely to exit the league without position controls and 43% more likely with position controls than non-Hispanic White players. In addition, we find that foreign born Hispanic players who come from elsewhere at 67% more likely to exit the league with and without position controls compared to non-Hispanic White players.

We find, however, that there is no economically or statistically significant difference in coefficients for Hispanics born in the United States or for Hispanics born in the US territory of Puerto Rico compared to non-Hispanic White players. Our results suggest that there is potentially exit discrimination present for foreign born players but not for players born within the United States or its territories. The result that foreign born players have higher likelihood of exit is also found in the NBA by Groothuis and Hill (2018) who find that foreign born players who did not attend college in the United States had higher probabilities of exiting the NBA. In addition, Depken, Ducking, and Groothuis (2017) find that European and Russian born players have a higher likelihood of exiting the NHL than players born in the United States or Canada. Our results coupled with their results suggest that foreign born players in American sports

leagues have shorter careers. These shorter careers may be caused by exit discrimination towards foreign players that push these players out of their sport or there may be pull factors that cause these players to return to their home countries.

Lastly, we find that the coefficient on Black is not statistically or economically significant in model (1) that excludes positions but becomes negative and significant in model (2) after we control for positions. A negative coefficient indicates that non-Hispanic Black players have longer careers than their performance indicates as compared to non-Hispanic White players. Converting the coefficient following the same method as above finds that Black players are 20% less likely to exit when position controls are included.

The coefficients on all the position dummy variables in model (2) are all positive and economically and statistically significant as compared to the catcher position. This indicates that first basemen, second basemen, third basemen, shortstops, outfielders, and utility infielders are all more likely to exit in comparison to catchers. Converting the coefficients to a percentage finds that first basemen have a 105% and third basemen a 74% higher hazard rate then catchers. Outfielder and second basemen have the highest hazard rate of 136% compared to catchers while shortstops have the lowest rate in the infield at 62% and utility infielders have the lowest of all positions at 38% as compared to catchers.

Considering that of all the positions outfielders have the highest hazard rates and that non-Hispanic Black players are most likely to play outfield indicates that even if Black players have a lower probability of exit after controlling for position, the fact that roughly two-thirds of the Blacks in the data play in the outfield leads us to question if Blacks are funneled to a position that has shorter careers than other positions. To explore occupational stacking or discrimination we estimate a multinomial logit on position played in the next section.

Table 4: Semiparametric Analysis of Career Duration

Variables	(1) Exit	(2) Exit
Black Non-Hispanic	-0.001 (0.079)	-0.221** (0.084)
Hispanic United States	-0.004 (0.158)	-0.092 (0.169)
Hispanic Puerto Rican	0.084 (0.141)	0.096 (0.142)
Hispanic Dominican	0.378** (0.134)	0.361** (0.134)
Hispanic Other	0.513** (0.128)	0.509** (0.129)
Height	0.021 (0.020)	0.017 (0.020)
Weight	0.001 (0.003)	0.001 (0.003)
Age	0.161** (.017)	0.177** (.017)
Slugging Average	-1.485** (0.537)	-2.067** (0.554)
Runs Batted in per game	-1.485** (0.349)	-1.243** (0.352)
Walks per game	-0.653** (0.307)	-0.699** (0.317)
Strikeouts per game	0.854** (0.181)	0.857** (0.185)
Stolen Bases per game	-3.018** (0.704)	-3.923** (0.750)
Games	-0.013** (0.001)	-0.013** (0.001)
1st Base	-	0.720** (0.143)
2nd Base	-	0.861** (0.150)
3rd Base	-	0.557** (0.123)
Shortstop	-	0.480** (0.144)
Outfield	-	0.856** (0.101)
Utility	-	0.302** (0.125)
Constant	-5.352** (1.372)	-5.909** (1.410)
Wald Pseudo R^2	803.29** 0.151	853.77** 0.162

*significant at the 10% level, **significant at the 5% level, clustered standard errors in parentheses. Catcher and non-Hispanic White reference categories.

Multinomial Logit Estimates of Positional Segregation

To explore positional discrimination, we utilize multinomial logit estimates similar to Sack et al. (2005) who used cross sectional data from the 1998 season. In our case we use panel data from 1990 to 2004 to estimate a multinomial logit model that uses catchers as the reference category. We chose catcher as the reference category because it has the highest percentage of non-Hispanic White players of all the positions.

In table 5, we present the multinomial logit estimates with catcher as the reference category. The positive and statistically significant coefficient on Black for all positions indicate that Blacks are more likely to play each of the positions listed in the table than they are to play catcher. Focusing on the magnitudes of the coefficients we find that the coefficients for Black players are the largest of all racial or ethnic groups. To provide insights we convert the coefficients to odds ratios and find that the odds of a Black player being a first basemen is 9 to 1 over being a catcher, being a second baseman 7 to 1, being a third baseman 4 to 1, being a shortstop 6 to 1, and being a utility infielder being 5 to 1. The biggest odds ratio for a Black player is for being an outfielder. The odds of a Black player being an outfielder is 29 to 1 over being a catcher. These results show that Black players are stacked in the outfield and seldom catchers even after controlling for size and performance statistics.

It is not clear if the sorting of Black players in Major League baseball to the outfield is due to higher skill and development cost (Medoff 1977, Medoff 1986, and Verduci 2003), the belief by individuals in charge that Blacks are incapable of performing at thinking positions (Smith and Leonard 1997), their heroes playing in the outfield and influencing them to play in the outfield (Smith and Leonard 1997), or Blacks have talent that is most suitable for playing in the outfield (Margolis and Piliavin 1999). All of these options are consistent with the outcome

that Blacks are more likely to play in the outfield although some may seem more viable than others.

Table 5: Multinomial Logit
Catcher Reference Category

Variables	1st Base	2nd Base	3rd Base	Shortstop	Outfield	Utility
Black	2.179**	1.979**	1.257**	1.892**	3.383**	1.641**
	(0.235)	(0.241)	(0.246)	(0.257)	(0.214)	(0.252)
Hispanic US born	0.819**	0.394	0.915**	1.718**	1.441**	0.687*
	(0.326)	(0.360)	(0.305)	(0.316)	(0.275)	(0.333)
Hispanic Puerto Rican	-2.921**	0.125	-1.753**	0.456**	0.062	-0.511**
	(0.595)	(0.202)	(0.290)	(0.208)	(0.160)	(0.241)
Hispanic Dominican	0.874**	0.284	0.177	1.021**	0.929**	0.647**
	(0.271)	(0.238)	(0.243)	(0.227)	(0.204)	(0.228)
Hispanic Other	0.718**	-0.854**	-0.540**	0.477*	0.359*	-0.706**
	(0.209)	(0.256)	(0.217)	(0.215)	(0.177)	(0.228)
Height	0.301**	-0.157**	0.029	0.227**	0.126**	0.064
	(0.036)	(0.037)	(0.034)	(0.038)	(0.028)	(0.037)
Weight	0.014**	-0.078**	-0.017**	-0.108**	-0.024**	-0.077**
	(0.004)	(0.005)	(0.004)	(0.006)	(0.003)	(0.005)
Age	-0.019	-0.234**	-0.080**	-0.298**	-0.073**	-0.001
	(0.028)	(0.030)	(0.026)	(0.031)	(0.021)	(0.027)
Experience	-0.040	0.185**	0.051*	0.314**	0.040	-0.039
	(0.028)	(0.030)	(0.026)	(0.032)	(0.022)	(0.028)
Slugging Average	4.563**	-3.689	1.792*	-1.289	7.694**	4.187**
	(0.949)	(1.03)	(0.910)	(1.096)	(0.762)	(0.910)
Runs Batted in per game	-0.645	-3.528**	-0.134	-2.427**	-3.624**	-4.708**
	(0.624)	(0.612)	(0.516)	(0.632)	(0.435)	(0.608)
Walks per game	3.453**	0.823	0.727	-1.293**	0.691	-0.834
	(0.459)	(0.489)	(0.424)	(0.527)	(0.365)	(0.523)
Strikeouts per game	-1.391**	0.322	0.660**	0.741**	0.883**	0.321
	(0.300)	(0.323)	(0.277)	(0.329)	(0.236)	(0.330)
Stolen Bases per game	-8.592**	26.623**	13.986**	21.714**	27.675**	20.543**
	(2.274)	(1.762)	(1.902)	(1.823)	(1.696)	(1.867)
Games	0.007**	0.009**	0.009**	0.016**	0.003*	0.005**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
Constant	-27.256**	-29.067**	0.751	9.076**	-5.751**	8.579**
	(2.519)	(2.550)	(2.299)	(2.625)	(1.887)	(2.456)

Log likelihood= -9746.39, Pseudo R^2 = 0.238

When focusing on the magnitudes of the various Hispanic groups, we find that for all groups the magnitudes are much smaller than for Black players. The one consistent result is that for all Hispanic groups they are more likely to be shortstops than catchers. The odds are 2 to 1 for Hispanics born in the United States, Puerto Rico or elsewhere in Latin America and 3 to 1 if

born in the Dominican Republic. In the case of the sorting of Hispanic players in Major League baseball to shortstop, a possibility is that they received training in their home country that allow them to be successful at shortstop.

When focusing on outfielders, we find that the odds are 2 to 1 over being a catcher for Hispanics born in the United States and the Dominican Republic while this is not true for other Hispanic groups. Lastly, we find that Hispanic players who were born in Puerto Rico are 19 to 1 more likely to be a catcher than a first baseman and 6 to 1 to be a catcher than a third baseman. Suggesting that Puerto Ricans are stacked in the more skilled positions of catcher and shortstop.

Overall, our results suggest that stacking occurs most for Black players. They are most likely to be placed in the least skilled positions while Hispanic players are more likely to be stacked into the more skilled positions of shortstop and in the case of Puerto Ricans both shortstop and catcher. These results from 1990s data are consistent with the conjecture that non-Hispanic Black players had less access to training in their youth while Hispanic players particularly from the Dominican Republic and Puerto Rico had more access to training. Both the lack of access and the greater access to training is consistent with the decrease of non-Hispanic Black players in the league and increase in Hispanic players in the league that occurred from the early 2000s until today.

V. Conclusion

We examine Major league baseball data from 1990 to 2004 and find evidence that foreign born Hispanics have shorter careers than non-Hispanic Whites after controlling for performance due to their higher likelihood of exit in comparison to non-Hispanic Whites. We also find that non-Hispanic Blacks are less likely to exit than non-Hispanic Whites after controlling for

position, but the position that most non-Hispanic Blacks in our data play has a higher likelihood of exit than other positions.

Even though we find evidence of positional segregation for non-Hispanic Blacks and Hispanics, it is not clear that it is due to discrimination. Future research can test to see if programs that are designed to improve the baseball skills of non-Hispanic Blacks decrease the likelihood that Blacks play in the outfield, and increase the likelihood that blacks play catcher and shortstop.

References

Becker, G. The Economics of Discrimination. 2nd edition. Chicago: University of Chicago Press. 1971.

Berger, M. and Black, D. The duration of Medicaid spells: An analysis using flow and stock samples. The Review of Economics and Statistics. 1998; 80:667-675.

Berger, M., Black, D. and Scott, F. Is there job lock? Southern Economic Journal 2004; 70:953-976.

Blinder, A. Wage discrimination: reduce form and structural estimates. Journal of Human Resources. 1973; 8:436–455.

Burnett, N. and Van Scyoc, L. Compensation discrimination for wide receivers: applying Quantile regression to the National Football League. Journal of Reviews on Global Economics. 2013; 2:433–441.

Burnett, N. and Van Scyoc, L. Compensation Discrimination for Defensive Players: Applying Quantile Regression to the National Football League Market for Linebackers and Offensive Linemen. Journal of Sports Economics. 2015; 16(4):375–389.

Cox, D. R. Regression models with life tables. Journal of Royal Statistical Society 1972; 34:187–220

Depken, Craig, Johnny Ducking and Peter A. Groothuis, "Career Duration in the NHL: Pushing and Pulling on Europeans?" Applied Economics. 2017; 40(59):5923–5934.

Ducking, J., Groothuis, P. and Hill, J. R. Compensation discrimination in the NFL: an analysis of career earnings. Applied Economic Letters. 2014; 21(10):679–682.

Ducking, J., Groothuis, P. and Hill, J. R. Exit Discrimination in the NFL: A Duration Analysis of Career Length. The Review of Black Political Economy. 2015; 42(3):285–299

Ducking, J., Groothuis, P. and Hill, J. R. Compensation Discrimination: an Analysis of Linebackers, Defensive Linemen, and Defensive Backs in the National Football League. The Review of Black Political Economy. 2017; 44:251–260.

Gius, M. and Johnson, D. An empirical investigation of wage discrimination in professional basketball. Applied Economics Letters. 1998; 5(11):703–705.

Gius, M. and Johnson, D. Race and compensation in professional football. Applied Economic Letters, 2000; 7(2):73–75.

Groothuis, P.A. and Hill, J.R. Exit discrimination in the NBA: A duration analysis of career length using flow and stock samples. Economic Inquiry. 2004; 42:341–349.

Groothuis, P.A. and Hill, J.R. Exit Discrimination in Major League Baseball: 1990–2004. Southern Economic Journal. 2008; 75(2):574–349.

Groothuis, P.A. and Hill, J.R. Pay discrimination, exit discrimination or both? Another look at an old issue using NBA data. Journal of Sports Economics. 2013; 14(2):171–185.

Groothuis, Peter A. and James Richard Hill, "Career Duration in the NBA: Do Foreigner Players Exit Early?" Journal of Sports Economics, Vol. 19 No. 6, August 2018.

Hill, J.R. Pay discrimination in the NBA revisited. Quarterly Journal of Business and Economics, 2004; 43(1/2), 81–92.

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

Holmes, P. New Evidence of Salary Discrimination in Major League Baseball. Labour Economics, 2011; 18(3):320-331
Jiobu, R. Racial inequality in a public arena: The case of professional baseball. Social Forces. 1988; 67:524-534.
Johnson, C. and Minuci, E. Wage discrimination in the NBA: Evidence using free agent signings. Southern Economic Journal. 2020; 87(2):517-539.
Kahn, L.M. Discrimination in professional sports: a survey of the literature. Industrial and Labor Relations Review. 1991; 44:395–418.
Kahn, L. M. The effects of race on professional football players' compensation. Industrial and Labor Relations Review. 1992; 45:295–310.
Kahn, L. M. The Sports Business as a Labor Market Laboratory. Journal of Economic Perspectives, 2000; 14(3):75-94.
Kahn, L.M. and Shah, M. Race, compensation and contract length in the NBA: 2001–2002. Industrial Relations: A Journal of Economy and Society. 2005; 44(3):444–462.
Kahn, L.M. and Sherer, P.D. Racial differences in professional basketball players' compensation. Journal of Labor Economics. 1988; 6(1):40–61.
Keefer, Q. A. W. Compensation discrimination for defensive players: applying Quantile regression to the National Football Market for linebackers. Journal of Sports Economics. 2013; 14(1):23–44.
Margolis, B. and Piliavin, J.A. Stacking in Major League Baseball: A Multivariate Analysis. Sociology of Sport Journal. 1999; 16(1):16-34.
Medoff, M. Positional segregation and professional baseball. International Review for the Sociology of Sport. 1977; 12(1):49-56.
Medoff, M. Positional segregation and the economic hypothesis. Sociology of Sport Journal. 1986; 3:297-304.
Mogull R. Football salaries and race: some empirical evidence. Industrial Relations. 1973; 12:109–112.
Mogull, R. Salary Discrimination in Major League Baseball. The Review of Black Political Economy. 1975; 5(3):269-279.
Oaxaca R. Male-female wage differentials in urban labor markets. International Economic Review. 1973; 14:693–709.
Palmer, M. and King, R. Has Salary Discrimination Really Disappeared from Major League Baseball? Eastern Economic Journal. 2006; 32(2):285-297
Sack, L., Singh, P. and Thiel, R. Occupational Segregation on the Playing Field: The Case of Major League Baseball. Journal of Sport Management. 2005; 19(3):300-318
Scully, G. W. Economic discrimination in professional sports. Law and Contemporary Problems. 1973; 38(1):67–84.
Simmons, R. Professional Labor Markets in the Journal of Sports Economics. Journal of Sports Economics. November 2021. doi:10.1177/15270025211051062
Smith, E. and Leonard, W. M. Twenty-five years of stacking research in Major League Baseball: An attempt at explaining this re-occurring phenomenon. Sociological Focus, 1997; 30(4):321-331.

Verducci, T. Blackout: The Black baseball player is vanishing:
Does he have a future? Sports Illustrated. 2003, July 7. Retrieved from:
https://vault.si.com/vault/2003/07/07/Blackout-the-Black-baseball-player-is-vanishing-does-he-have-a-future