Appalachian State University

Department of Economics Working Paper

Number 10-04 | March 2010

Race and Survival Bias in NBA Data

Peter A. Groothuis
Appalachian State University

James Richard Hill
Central Michigan University

Department of Economics
Appalachian State University
Boone, NC 28608
Phone: (828) 262-6123
Fax: (828) 262-6105
www.business.appstate.edu/economics

Race and Survival Bias in NBA Data

by

Peter A Groothuis
Professor
Appalachian State University
Boone, NC

and

James Richard Hill
Professor
Central Michigan University
Mt Pleasant, MI

Spring 2010

Abstract: Cross sectional employment data is not random. Workers who survive to a longer level of tenure tend to have a higher level of productivity than those who exit earlier. Wage equations that use cross sectional data could be biased from the over sampling of high productive workers at long levels of tenure. The survival bias that arises in cross sectional data could possibly bias the coefficients in wage equations. This could lead to false positive conclusions concerning the presence of pay discrimination. Using 1989-2008 NBA data we explore the extent of survival bias in wage regressions in a setting in which worker productivity is extremely well documented through a variety of statistical measures. We then examined whether the survival bias affects the conclusions concerning racial pay discrimination.

Keywords: NBA, survival bias, pay discrimination

JEL classification codes: J7, J4

I. Introduction

Findings of discrimination in labor markets tend to garner much attention. This was particularly true when Kahn and Sherer (1988) proclaimed there was pay discrimination in the NBA for black players despite a roughly similar prevailing average salary. The magnitude of the pay discrimination was said to be in the 20% range. This stirred quite a discussion in a sport where about 75%-80% of the players were black. A summary of the findings even made it to the popular press. Proclamations of racial discrimination always elicit notoriety. Findings of no discrimination do not procure the same response. Therefore it is important that any positive findings of racially unequal treatment be particularly robust.

In addition to findings of wage discrimination in the NBA there have also been findings of exit discrimination. Hoang and Rascher (1999) found evidence that white players were employed longer in the NBA than comparable black players. This form of discrimination could be responsible for an even greater shortfall in pay for black athletes than the more commonly researched wage discrimination.

From research on exit discrimination it is clear that individuals with greater ability have a higher survival rate. This leads to an over-representation of high ability workers at high levels of experience. This suggests that the coefficients of explanatory variables, including race, may be biased because in a cross sectional data set the observations are not random.

In this article we explore a modified-Heckman correction technique to correct for potential survival bias in wage equations using data from the 1989-2008 time period. This technique provides a logical extension of previous work on pay discrimination and exit discrimination by incorporating the results of exit discrimination into the wage equations directly. In the next section we provide a review of the previous literature on pay discrimination and exit discrimination in the NBA. Then we explain the duration model of career length utilized in our analysis and present the empirical results of this model. In section four, we discuss how the inverse-Mills ratio can be used to correct for survival bias in wage equations. We then present the regression results for our wage equations using pooled OLS estimates using both a survival-bias corrected model and a standard model. Lastly, we conclude with a discussion of the results concerning racial pay discrimination in the NBA. We also suggest possible extensions of the survival bias technique for future research.

II. Literature Review

Following the article of Kahn and Sherer (1988) which used salary data from the 1985-86 season, articles by Brown, Spiro, and Keenan (1991) and Koch and Vander Hill (1988) using 1984-1985 season salary estimated pay premiums for white players of 15% and 9% respectively. All three studies cite fan discrimination as the driving force behind pay discrimination.

Articles by Dey (1997), Gius and Johnson (1998), and Bodvarsson and Brastow (1999) using data from the late 1980's and 1990's failed to find racial wage discrimination. Hamilton (1997) found evidence of racial pay differences only at the upper end of the 1994-1995 season's salary distribution. Bodvarsson and Brastow (1999) suggested the disappearance of the pay discrimination in the NBA was the result of a decrease in owner monopsony power due to the negotiation of a new NBA Collective Bargaining Agreement in 1988 combined with the addition of four new teams. They also provide empirical evidence that at least part of the racial salary gap was the result of owner and manager discrimination, not just fan discrimination. Except for the article by Dey (1997), all the research in this area used single season salary data for model estimation. Dey (1997) used pooled data from five seasons in some of his empirical work.

Hill (2004) used an eleven year panel data set beginning with 1990 salaries and ending with 2000 salaries. While the coefficients for the dummy variable for white players was positive and significant in 1990 and 1991individual year regression equations, the positive and significant finding in the overall pooled regression disappeared when height was added as an explanatory variable. Using data from the same year as Kahn and Sherer (1988) Hill also finds that the significance of race as a determinant of pay wanes when height is added to the model. In an attempt to revive this decades old controversy Kahn and Shah (2005) claim;

"For players who were neither free agents nor on rookie scale contracts, there were large statistically significant *ceteris paribus* nonwhite shortfalls in salary, total compensation, and contract duration." Only 21 white players and 75 nonwhite players fell into this category.

Hoang and Rascher (1999) define exit discrimination as "the involuntary dismissal of workers based on the preferences of employers, coworkers, or customers." They concluded that career length for black players in the NBA were lower than their white counterparts, ceteris paribus. Hoang and Rascher (1999) calculate that this form of discrimination led to almost a two and a half times greater decrease in black career pay compared to the more heavily analyzed form of pay discrimination. Groothuis and Hill (2004) focusing on career duration in the National Basketball Association (NBA) found that performance and weight of the basketball player determined career length. They found that the race of the player did not matter.

III. Duration Analysis

To test for survival effects we estimate semi-parametric hazard functions following Berger and Black (1998), Berger, Black, and Scott (2004), and Groothuis and Hill (2004); since our data is at the season level we calculate our hazard model as a discrete random variable. As with Berger, Black, and Scott (2004), we model the durations of a single spell and assume a homogeneous environment so that the length of the spell is uncorrelated with the calendar time in which the spell begins. This assumption lets us treat all the players' tenure as the same regardless of when it occurred in the panel study. For instance, all fourth year players are considered to have the same base line hazard regardless of calendar time so a fourth year player in 1990 has the same baseline hazard as a fourth year player in 1997.

To understand how stock data influences a likelihood function we follow the notation of Berger, Black, and Scott (2004). 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

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

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

$$L(\beta) = \prod_{i=1}^{n} f(t_i, x_i, \beta) = \prod_{i=1}^{n}\left(\prod_{j=1}^{t_i-1}[1 - h(j, x_i, \beta)]\, h(t_i, x_i, \beta)\right). \qquad (2)$$

Often it is not possible to observe all careers until they end, hence careers are often right-censored. 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).

To introduce stock sampling, 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. Of course, 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 1989 and ends his career after 2008, an unlikely event. Taking into account all four sets: A, B, C, and D the likelihood function becomes

$$L(\beta) = \prod_{i\in A}\left(\prod_{j=1}^{t_i-1}[1-h(j,x_i,\beta)]\,h(t_i,x_i,\beta)\right) \times \prod_{i\in B}\left(\prod_{j=1}^{t_i-1}[1-h(j,x_i,\beta)]\right) \times \prod_{i\in C}\left(\prod_{j=r_i}^{t_i-1}[1-h(j,x_i,\beta)]\right)h(t_i,x_i,\beta) \times \prod_{i\in D}\left(\prod_{j=r_i}^{t_i-1}[1-h(j,x_i,\beta)]\right) \quad (3)$$

In equation (3) 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).

Thus 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 NBA given that the NBA 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.   For our purposes we choose the probit model. We choose the probit model over the more traditional logit model because the probit model provides the benefit of the ability to calculate the inverse-Mills ratio. We will review the implication of the inverse-Mills ratio in the next section.

The intuition behind the probit model for the hazard function is relatively simple. For each year during the survey in which the player is in NBA, the player either comes back for another season or ends his career.   If the career ends, the dependent variable takes on a value of one; otherwise, the dependent variable is zero. The player remains in the panel until the player exits NBA or the panel ends. If the panel ends, we say the worker's spell is right-hand censored. Thus a player who begins his NBA career during the panel and plays for 6 years will enter the data set 6 times: the value of his dependent variable will be zero for the first 5 years (tenure one through five) and be equal to one for the sixth year.

To illustrate a stock sample consider another player who enters the panel with 7 years of NBA job tenure prior to 1989 the first year of the panel, then plays for an additional 3 years for a 10 year career. For this player we ignore his first 7 years of tenure because he is left-hand censored. As the equation of the likelihood function with stock data indicates, the duration of a NBA career prior to the beginning of the panel makes no contribution to the value of the likelihood function. Therefore only years 8 through 10 will enter the data set with the dependent variable taking on the value zero for years 8 and 9 and in the 10^{th} year it takes on a value of one with this player appearing in the data set a total of 3 times. Note for all players who are right-hand censored, we do not know when their career ends so their dependent variables are always coded as zero.

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 (3) can be very large. In our study it would require 21 dummy variables. We also run into problems with the dummy variable 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 4th order polynomial of the players' tenure in NBA, which reduces the number of parameters to be estimated from 21 to 4. Thus, the hazard function becomes

$$\Pr(t, x\beta) = \Pr(\phi(t) + x\beta), \qquad (4)$$

where $\phi(t)$ is a 4th order polynomial of the player's tenure in NBA. The 4th order polynomial therefore includes tenure to the first, second, third, fourth and fifth powers. Once again, we choose the Taylor series approximation technique over using tenure dummies due to the small number of observations for high tenures.[1]

In table 2 we report the results of equation 6 for NBA players. Independent variables included in the model are games played during the previous season, a measure of player performance, draft number, and biographical information.[2] The measure of player performance is called the *efficiency formula*. As reported by NBA.com, this index is calculated per game as: (points + rebounds + assists + steals + blocks) – ((field goals attempted – field goals made) + (free throws attempted - free throws made) + turnovers)). Biographical data included weight and

[1] When higher order polynomials of the fifth and sixth power are included results do not change suggesting that a fourth order polynomial is flexible enough to capture the influence of the base line hazard.

[2] Annual statistics were obtained from Doug's MLB and NBA Stat Home: http://www.dougstats.com/. Biographical information on each player came from a variety of sources including the third edition of the NBA Encyclopedia, various editions of The Sporting News Official NBA Player Register, Wikipedia, and/or NBA.com.

race (a dummy variable equal to one for a white player). Draft number ranged in value from one for the first player selected in the draft to 60 for the last player selected in the second round.[3] We find that efficiency, games played in the previous season, weight, and higher draft position all significantly lower the probability of exiting a career. Race does not have a significant impact on survival. These results show that performance lengthens the careers of NBA players suggesting that higher performance players are over-represented latter in careers. In the next section we explore if survival bias effects wage equations.

IV. Duration effects and Sample Selection Bias

The Heckman (1979) procedure has long been used by economist to control for self selection in a sample. In our case, we suggest that performance increases the likelihood of survival in the NBA. We explore if a modified Heckman procedure controls for survival bias in wage equations. Following Green's (1998) notation consider the basic sample selection model where,

$$d_i^* = z'\gamma + \upsilon \text{ and } d_i = 1 \text{ if } d_i^* > 0;\ d_i = 0 \text{ otherwise,} \qquad (5)$$

$$y_i = x'\beta + \varepsilon_i \qquad (6),$$

and y is only observed if d_i is equal to one. Suppose as well that υ and ε have a bivariate normal distribution with zero means and a correlation of ρ. Then

$$E(y|\ y \text{ is observed}) = E(y|\ d^* > 0) \qquad (7)$$

$$= E(y|\ \upsilon > -z'\gamma) \qquad (8)$$

$$= x'\beta + E(\varepsilon|\ \upsilon > -z'\gamma) \qquad (9)$$

$$= x'\beta + \rho\sigma_\varepsilon\lambda(\alpha_\upsilon) \qquad (10)$$

[3]Since 2005 there are 30 teams in the league. Beginning with the 1989 draft there were only two rounds. There were three rounds in the 1988 draft. There were 7 rounds from 1985-1987; there were 10 rounds from 1974-1984. Draft numbers in these seasons had much higher ranges in value. For undrafted players in seasons with only two rounds players were assigned a draft number of 65.

or $= x'\beta + \beta_\lambda \lambda(\alpha_\upsilon)$ (11)

Where $\lambda(\alpha_\upsilon)$ is the inverse Mill's ratio. The overall equation of interest becomes

$$y_i|d^*>0 \,|= x'\beta + \rho\sigma_\varepsilon\lambda(\alpha_\upsilon) + \varepsilon_i, \quad (12)$$

Thus the marginal effect of a regressor in the observed sample consists of two parts

$$\frac{\partial\, E(y|\, y \text{ is observed})}{\partial x_k} = \beta_k + \gamma_k\, (\rho\sigma_\varepsilon\ /\ \sigma_\upsilon)\lambda(\alpha_\upsilon) \quad (13)$$

where the first part β is the marginal effect of the observed equation and γ is the influence the variable has on the likelihood of being observed. Heckman (1979) recommends a two step procedure: first estimate a probit on the selection equation to obtain estimates of γ. Then estimate the OLS regression along with the inverse Mill's ratio to estimate β and β_λ.

In our case the selection equation comes from equation 5 from the above career duration model. The probit estimates the probability of exiting from a career so instead of estimating the probability of being in the sample, as in Heckman's approach, we estimate the probability of exiting the sample. Equation 6 in our analysis is a cross sectional wage equation where the data on workers are selected into the sample with higher skilled workers remaining in the sample while low skill workers are more likely to leave.

In table 3, we report the results of the contemporaneous log wage equation for NBA players. Salary data was not available for the 1989 season so the regression covers the 1990 through 2008 seasons.[4] Independent variables include the same list used for the probit model of Table 2 with the exception of weight. Following the findings of Hill (2004) height is substituted for weight in the wage equation.[5] Years and years squared are inserted as explanatory variable to capture the traditional shape of the age earnings profile. Dummy variables are inserted for each

[4] Salary figures were obtained from Patricia Bender's website: http://www.eskimo.com/~pbender/index.html
[5] Height and weight have a high degree of correlation, around 70%. Since white players in the league are taller on average we have chosen to use height in the wage equation.

year except 2008 to control for inflationary factors in pooling yearly data. The inverse Mill's ratio is used as an additional independent variable to correct for survival bias in models II and IV listed in Table 3. Models I and II omit the height variable to determine if the race dummy variable is picking up correlation from height.

The results of the log wage regressions in Table 3 are both predictable and surprising at the same time. First the expected results: in all four models the coefficients of years are positive and highly significant; the coefficients of years squared are negative and highly significant. The shape of the age/earnings profile follows the classic pattern found throughout labor markets in general. Wages are also negatively and very significantly correlated with draft number; earlier picks in the draft earn more money. The number of games played in the previous season has a positive and significant impact on wages. Player performances, as measured by the efficiency variable, and height have a positive and significant effect on pay. The dummy variables for each year follow the expected pattern: each year has a negative and significant effect on wages; only the coefficient of the dummy variable for 2007 is not significant. The absolute value of the coefficient for each year's dummy variable continually decreases in value and level of significance. The only exception to this pattern is the 1998-1999 time period when the league experienced a lockout before signing a new Collective Bargaining Agreement in 1999.

The findings concerning the dummy variable for white players are surprising. In Model I and II when the height variable is left out of the equation the coefficient for White is negative but insignificant. When height is included in the model the coefficient for White is negative and significant. The coefficient of the Inverse Mills Ratio is negative and significant in both Models II and IV. Survival bias appears to exist in the determination of wages but surprisingly the correction for survival bias does not significantly change the coefficients of any variables in the model.

V. Conclusions, Implications, and Suggestions for Future Research

The empirical results of this study would lead one to conclude that there is either no pay discrimination in the NBA or that there is pay discrimination against white players. This would certainly represent a complete turnabout from the earlier research findings of Kahn and Sherer (1988) and others. In a twisted sort of way one could argue that there was pay discrimination against the minority group since white players are underrepresented in the league. Yet in a broader context it would appear much more honest to suggest that the determination of wage discrimination using the residual methodology is tricky at best. Results do not appear to be robust.

The failure of the correction for survival to highlight bias in the coefficients of the model is interesting. On the one hand researchers using cross sectional data might feel satisfied that the lack of bias found in our research might mean such a bias does not exist. Caution should be exercised in this judgement. Sports wages are extremely deterministic compared to other market structures in which good performance variables are lacking. The technique introduced here to correct for survival bias needs to be tried in other labor market settings. This could prove a fruitful area for future research.

TABLE 1
Means

Variable	Means
Career length	5.83(seasons)
Height	79.15 (inches)
Weight	221.13 (lbs.)
Games per season	60.31
Efficiency	9.54
Draft Number	26.27
Number observations	6530
White	21.45% of observations

TABLE 2

NBA Career Duration Semi-parametric Analysis
1989-2008

Variable	Probit Model
Constant	-.2779 (1.140)
White	.0381 (.642)
Weight	-.0021 (2.364)
Games	-.0252 (21.632)
Efficiency	-.0023 (7.039)
Draft Number	.0082 (8.155)
Chi-squared	841.92107
Number observations	6530

--First through fourth order tenure polynomials are also included to provide for general functional form of baseline hazard. They are jointly significant.
--The numbers in parentheses are absolute value of t-ratios.

Table 3: Log Wage Regressions: 1990-2008

Variable	I	II	III	IV
Constant	13.6901 (258.414)	13.6488 (262.676)	11.7687 (58.539)	11.7702 (59.759)
Years	.3150 (40.477)	.3157 (41.413)	.3182 (41.161)	.3189 (42.101)
Years Squared	-.0155 (30.592)	-.0156 (31.306)	-.0158 (31.315)	-.0159 (32.031)
Draft Number	-.0163 (41.086)	-.0164 (41.992)	-.0156 (38.831)	-.0156 (39.729)
White	-.0312 (1.454)	-.0319 (1.518)	-.0838 (3.817)	-.0834 (3.875)
Height			.0240 (9.902)	.0235 (9.883)
Efficiency	.0006 (3.980)	.0006 (4.051)	.0006 (4.083)	.0006 (4.154)
Games	.0098 (21.330)	.0097 (21.548)	.0100 (21.906)	.0099 (22.125)
Inverse Mills Ratio		-.2834 (16.546)		-.2812 (16.534)
Year = 1990	-1.5025 (26.151)	-1.4581 (25.877)	-1.5088 (26.454)	-1.4646 (26.182)
Year = 1991	-1.3848 (25.445)	-1.3453 (25.207)	-1.3891 (25.714)	-1.3498 (25.478)
Year = 1992	-1.2726 (23.363)	-1.2213 (22.848)	-1.2764 (23.607)	-1.2255 (23.095)
Year = 1993	-1.1607 (21.593)	-1.1139 (21.123)	-1.1680 (21.888)	-1.1214 (21.421)
Year = 1994	-1.0798 (19.924)	-1.0413 (19.595)	-1.0866 (20.198)	-1.0483 (19.871)
Year = 1995	-.9455 (17.922)	-.8839 (17.057)	-.9567 (18.264)	-.8953 (17.400)
Year = 1996	-.9004 (17.0450	-.8517 (16.431)	-.9068 (17.292)	-.8583 (16.681)
Year = 1997	-.7682 (14.345)	-.7441 (14.178)	-.7750 (14.578)	-.7510 (14.413)
Year = 1998	-.4154 (7.605)	-.4065 (7.598)	-.4170 (7.692)	-.4082 (7.686)
Year = 1999	-.4673 (8.817)	-.4289 (8.252)	-.4726 (8.983)	-.4344 (8.420)
Year = 2000	-.3341 (6.274)	-.2852 (5.460)	-.3369 (6.374)	-.2885 (5.562)
Year = 2001	-.3076 (5.751)	-.2457 (4.676)	-.3188 (6.004)	-.2572 (4.930)
Year = 2002	-.2593 (4.763)	-.2152 (4.031)	-.2657 (4.918)	-.2219 (4.186)
Year = 2003	-.2858 (5.289)	-.2394 (4.515)	-.2931 (5.463)	-.2468 (4.689)
Year = 2004	-.2321 (4.352)	-.1759 (3.359)	-.2350 (4.439)	-.1792 (3.447)
Year = 2005	-.1368 (2.570)	-.0848 (1.623)	-.1419 (2.686)	-.0902 (1.739)
Year = 2006	-.1248 (2.354)	-.0521 (1.00)	-.1233 (2.344)	-.0512 (.991)
Year = 2007	-.0842 (1.548	-.0015 (.028)	-.0844 (1.563)	-..23 (.043)
R^2	.5859	.6026	.5920	.6085
Number observations	6530	6530	6530	6530

--The numbers in parentheses are absolute value of t-ratios.

References

Berger, Mark C. and Dan A. Black. The duration of Medicaid spells: an analysis using flow and stock samples. The Review of Economics and Statistics 1999; 80 (4); 667-674.

Berger, Mark C., Dan A. Black, and Frank Scott. Is there job lock? Evidence from the pre-HIPAA era. Southern Economic Journal 2005; 70 (4); 953-976.

Bodvarsson, Orn, B. and Ramond T. Brastow. A test of employer discrimination in the NBA. Contemporary Economic Policy 1999; 17(2); 243-255.

Brown, E., Spiro, R., and Keenan, D. Wage and non-wage discrimination in professional basketball: do fans affect it? American Journal of Economics and Sociology 1991; 50; 333-345.

Burdekin, Richard C. K., and Todd L. Idson. Customer preferences, attendance and the racial structure of professional basketball teams. Applied Economics 1991; 23; 179-186.

Dey, Matthew S. Racial differences in National Basketball Association players' salaries: A new look. American Economist 1997; 41(Fall); 84-90.

Gius, Mark and Johnson, Donn. An empirical investigation of wage discrimination in professional basketball. Applied Economics Letters 1998; 5; 703-705.

Groothuis, Peter A. and J. Richard Hill. Exit discrimination in the NBA: a duration analysis of career length. Economic Inquiry 2004; 42 (2); 341-349.

Groothuis, Peter A. and J. Richard Hill. Exit discrimination in Major League Baseball: 1990-2004. Southern Economic Journal 2008; 75 (2); 574:590.

Hamilton, Barton H.. Racial discrimination and professional basketball salaries in the 1990's. Applied Economics 1997; 29; 287-296.

Heckman, James. Sample selection bias as a specification error. Econometrica 1979; 47; 153-61.

Hill, James Richard. Pay discrimination in the NBA revisited. Quarterly Journal of Business and Economics 2004; 43 (1&2); 81-92.

Hoang, Ha, and Dan Rascher. The NBA, exit discrimination, and career earnings. Industrial Relations 1999; 38; 69-91.

Johnson, Norris R., and Marple, David P. Racial Discrimination in Professional Basketball: An Empirical Test. Sociological Focus 1973; 6 (Fall); 6-18.

Kahn, Lawrence M. Discrimination in professional sports: A survey of the literature. Industrial and Labor Relations Review 1991; 44 (April); 395-418.

Kahn, Lawrence M. and Malav Shah. Race compensation and contract length in the NBA: 2001-2002. Industrial Relations 2005; 44 (July); 444-57.

Kahn, Lawrence M. and Peter D. Sherer. Racial differences in professional basketball players compensation. Journal of Labor Economics 1988; 6 (January); 40-61.

Koch, James V. and C. Warren Vander Hill. Is there discrimination in the 'Black Man's Game'? Social Science Quarterly1988; 69(1);83-94.