University of Connecticut

Department of Economics Working Paper Series

Minority Status and Managerial Survival in Major League Baseball

Brian Volz
University of Connecticut

Working Paper 2008-36

September 2008

---

341 Mansfield Road, Unit 1063
Storrs, CT 06269–1063
Phone: (860) 486–3022
Fax: (860) 486–4463
http://www.econ.uconn.edu/

Abstract

The effect of minority status on managerial survival in Major League Baseball is analyzed using survival time analysis and data envelopment analysis. Efficiency scores based on team performance and player salary data from 1985 to 2006 are computed and included as covariates in a survival time analysis. It is shown that when controlling for performance and personal characteristics minorities are on average 9.6 percentage points more likely to return the following season. Additionally, it is shown that winning percentage has no impact on managerial survival when efficiency is controlled for.

Journal of Economic Literature Classification: J71, L83, C41

Keywords: Baseball, Management, Race, Survival, DEA

Introduction

Due to the obsessively accurate and detailed record keeping of Major League Baseball and the popularity of the sport, there exists a large amount of literature on almost all aspects of the sport, including discrimination and managerial retention. However, previous research has failed to link the two subjects by including race as a factor when examining managerial retention in professional sports. The analysis presented in this paper looks to fill this gap in the literature by providing a thorough analysis of the effects of minority status on managerial survival in Major League Baseball.

On opening day of 2007 only six out of thirty Major League Baseball teams were lead by minority managers. This represents only 20 percent of major league managers. This low percentage of minority managers is very surprising given the great diversity of players in Major League Baseball. Approximately 42 percent of major league players are black, Hispanic, or Asian. One would expect a lower percentage of minority managers than players as many minority players do not possess the English speaking skills required to manage. However, a difference this large certainly deserves investigation. Even more surprising than this relatively low percentage of minority managers, is the fact that for the 2007 season there was only one black manager in the major leagues. This is the lowest number of black managers since Frank Robinson became the first black manager over 30 years ago in 1975. Not only is the number of black managers at an all time low, but the number of minority managers has decreased significantly in recent years. From 2002 to 2007, the percentage of minority managers in the major leagues was reduced significantly from its all-time high of 33 percent to 20 percent.

This paper examines whether this recent downward trend in minority managers is the result of minorities having a lower probability of survival than white managers. This question is analyzed by applying survival time analysis and data envelopment analysis to data on major league managers from 1985 through 2006. Through survival analysis, it can be established whether or not minority status has a negative impact on the likelihood of a manager returning for another season.

Previous Literature

The following studies are most directly related to survival analysis of major league managers. Scully (1994) uses survival analysis to show that managerial retention is influenced by managerial efficiency for the sports of baseball, basketball, and football. However, Scully (1994) does not include other covariates in addition to managerial efficiency. I improve upon his survival analysis by including covariates for various characteristics including race. I also consider a wider range of underlying distributions for the survival analysis in this paper. Scully (1994) uses a measurement of managerial efficiency which is based on a comparison of a manager's actual winning percentage to an estimate of their maximum possible winning percentage. These maximum winning percentages are estimated using ratios of a manager's runs or points to his opponent's runs or points. The efficiency scores used in my analysis differ greatly from that of Scully (1994) as I use a data envelopment analysis (DEA) approach. In this analysis efficiencies measure a manager's performance relative to the best performance possible with his team's given talent as measured by player salaries. Unlike Scully (1994), the efficiency measure presented here does not rely on team scoring statistics, which could be affected by the quality of management. The DEA approach presented here is similar to that of Einolf (2004). However, I choose a different measure of output and include a negative input for the level of competition a team faces.

Fizel and D'Itri (1997) use probit estimation to study the impact of organizational performance on manager succession in 147 college basketball teams from 1984 to 1991. They use data envelopment analysis to measure managerial efficiency based on team talent and the strength of the opposition. As with the other previous studies Fizel and D'Itri (1997) do not include race in their analysis. They find that when winning percentage is included managerial efficiency does not have a significant impact. Using survival time analysis instead of probit estimation I find the opposite to be true for Major League Baseball. The advantage of survival time analysis over probit is that it allows the effect of the covariates to vary depending on the year of tenure being considered. This is desirable in this context as there is no reason to assume the effects of race would be the same in a manager's first year as in their tenth year.

Survival analysis is used by Audas, Dobson, and Goddard (1999) to examine the effects of various performance and descriptive variables on the tenure of English soccer coaches using individual game data from 1972 to 1997. While Audas et al. (1999) do utilize the Cox proportional hazard model, which is the model of choice in this study, they also do not include race as a covariate. Additionally, they use the team's league standing when the manager started his tenure to capture differences in the talent available to the manager. The data envelopment analysis presented in this study captures differences in talent through the use of player salaries. This data envelopment analysis measure is more appropriate for professional baseball as managers are not in charge of acquiring and exchanging players as they traditionally are in European soccer.

Variable Selection

For this analysis the dependent variable of interest is whether or not a manager returns for another season. This analysis treats all terminations equally and does not differentiate between voluntary and involuntary terminations. It is not uncommon in professional baseball for a manager to be allowed to resign from his job in order to save his reputation. This makes distinguishing between voluntary and involuntary terminations extremely difficult. Therefore, the dependent variable in this analysis is equal to zero for any season after which the manager returns to manage the same team and one for any season after which the manager does not return to coach the same team.

Managers are included in the analysis so long as they managed more than five consecutive games in a given season. The goal of this restriction is to eliminate those managers who were not actually being considered for retention, such as those filling in for an absent manager or during the transition to a new manager. There is no specific number of games which signals a manager is being considered for retention but it is safe to say those managing less than six games are not being considered for retention. Raising this restriction to 50 games would only eliminate 5 out of 87 minority observations and does not affect the conclusions of this analysis. Therefore, the author chooses to only eliminate those who are most obviously not under consideration for retention without fear that this restriction is driving the results of the analysis.

The dependent variable or probability of returning in the next season is expected to be influenced by both the manager's performance and by individual characteristics of the manager and team. The first and most important job of a manager is to win games. Therefore, it is expected that some measure of team wins would have a significant positive impact on the probability that a manager returns the following season. The winning percentage for each team during the portion of the year that the manager was with that team is used to measure this ability to win. Winning percentage is chosen over total wins as it makes comparisons between managing spells of different lengths possible.

While regular season wins are important, the ultimate goal of a team is not to win regular season games but rather to win the World Series. Therefore, playoff wins are also expected to influence managerial survival. In order to measure a manager's playoff success, wins in the League Championship Series and World Series are also included in the survival analysis. Wins in the Division Series are not included in this analysis as the division series did not exist prior to the 1995 season.

While winning is the most important thing to baseball fans, owners must also consider profits. Due to these financial restrictions some managers may be given less talent to work with than others. When faced with one of these low budget rosters there is only so much a manager can do to make a team win. If a team has more talent than other teams and still loses it is likely that people will look at the management as a source of the problem. However, if a team with no talent loses it is unlikely that the manager will be blamed. Therefore, a more appropriate measure of manager performance may be how efficiently the manager transforms his given resources into wins. Efficiency scores for managers can be calculated using data envelopment analysis and then included in the survival analysis to control for managerial performance. The specifics of the data envelopment analysis are discussed in the next section.

It is obvious that managerial survival is not completely based on wins and league standing. In some cases the decision to retain or fire a manager may be based on personal characteristics and qualifications. One important characteristic is the experience of a manager. It is expected that years spent in baseball as a coach or player will increase a manager's baseball specific human capital and therefore make them a better coach. Such increases in on the field performance should be captured by the efficiency measure. However, experience may also improve skills not captured by the efficiency measure, such as the ability to communicate with the press and interact with front office personnel. For this reason it is expected that experience as a manager should increase the likelihood of retention beyond its contribution to efficiency. It is also expected that as experience increases its returns in terms of human capital may diminish as the probability of learning something new decreases with years spent around baseball. In order to capture the effects of this human capital on the probability of survival, games as a manager at the Major League level are included in the analysis. The square of this variable is also included to capture the expected diminishing effects.

Similarly to experience, one would expect that age also influences the probability of a manager returning for the next season. The amount of information and experiences a manager has been exposed to in and out of baseball increases with age. Any useful information or experience outside of managing games would not be captured by the experience variable. This would lead one to expect a positive relationship between age and retention. However, as age increases it is also more likely that a manager will choose to retire. It may also be expected that as managers' ages increase they are less able to relate to young players and will therefore be less effective. This increased likelihood of retirement and inability to relate to young players leads to a negative relationship between age and retention. To capture these effects and determine whether they are positive or negative managers' ages, in years, are calculated and included in the analysis. As with experience we would expect that the benefits of age would diminish with time and the likelihood of retirement would increase with age and so the square of age is also included in the analysis.

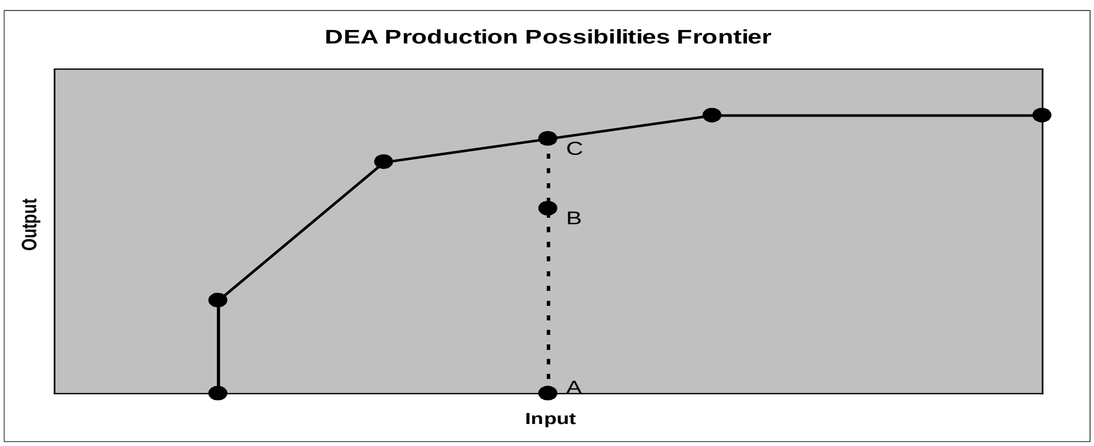
The personal characteristic of interest in this analysis is the manager's minority status. A variable which is equal to one for Black and Hispanic managers and zero otherwise is included to capture this effect. If minorities face discrimination in managerial retention decisions, one would expect to observe a negative relationship between this variable and the probability of survival.

Measuring Efficiency with Data Envelopment Analysis

In order to evaluate how efficiently a manager turns his team's potential into wins the method of data envelopment analysis can be applied. This application differs little from the use of data envelopment analysis in productivity analysis of firms. In productivity analysis firms take inputs and transform them into outputs. These outputs are compared to a constructed production possibilities frontier in order to determine how efficient the firm is. The analysis presented here is based on the output oriented technical efficiency method presented by Banker, Charnes, and Cooper (1984).

The production frontier is based on three assumptions about the production technology. The first assumption is that inputs are freely disposable. This means that if a certain level of inputs can produce some level of output, then a level of inputs which is greater in at least one dimension can also produce that level of output. The second assumption is that output is freely disposable. This implies that if a certain level of inputs can produce some level of output, then that same level of inputs can also produce any lower level of output. The third assumption is the convexity of the production possibilities set. Convexity implies that any linear combination of two feasible points is also feasible.

This production frontier can be shown graphically for the one input one output case. Given a set of observed points it is possible to show all points which are feasible. The convexity assumption means that all points which are linear combinations of the observed points are possible. Graphically, this is shown by connecting the data points to create a convex hull. The assumption of free disposability of inputs implies that all points to the right of the convex hull are feasible. Similarly the assumption of free disposability of output implies that all points below a feasible point are possible. When these assumptions are combined the result is a production possibility set as seen in the following figure.



In order to measure how efficiently a firm, or in this case manager, is producing, their actual output can be compared to the maximum feasible output given that level of inputs. This is referred to as the output oriented technical efficiency. This measure of efficiency is most appropriate in this analysis as managers have limited ability to choose their inputs as building the team is the job of the general manager. Therefore, the question of interest is how much more could the manager have produced given his inputs rather than how much could inputs have been reduced given the managers output. The technical efficiency is measured by the firms actual output divided by the maximum feasible output for that level of inputs. In the above chart, for observation B, this is the distance from A to B divided by the distance from A to C.

In the case of baseball managers, the output of interest is wins. However, due to the fact that not all managers manage an entire season and not all seasons have had the same number of games wins are not comparable from one observation to the next. This problem is avoided by measuring output in terms of winning percentage. For this data envelopment analysis playoff wins are not included as managers who are given teams of very low ability will likely not have the possibility of producing any playoff wins. Therefore, only regular season performance is included in this analysis. However, playoff wins are included in the survival analysis presented in the next section.

While the output of interest seems very obvious and easy to calculate the inputs are much more difficult. The inputs of interest are the players a manager is given. More specifically the inputs are the amount of talent which a manager has to work with. This talent can come in many forms, such as hitting, fielding, pitching, and speed. These talents are traditionally measure through statistics such as batting average, earned run average, and runs scored. However, in this context using measures of player performance is inappropriate. This is due to the fact that a player's individual performance over a season is likely to be influenced by the quality of the manager. A good manager may cause a pitcher to have a lower ERA than a bad manager. If input is measured in terms of ERA then the measure will always be overstating the level of input for good managers and understating the level of input for bad managers. Simply using performance statistics from prior seasons would not eliminate this problem for those players who have only played under one manager. Additionally, past statistics would not exist for any players who are in their first season at the major league level.

What is needed is a measure of player potential independent of the influence of the manager. It is reasonable to assume that teams pay players based on how they expect them to perform. This level of pay is determined before the player actually plays that season and therefore should not be influenced by the manager's performance that season. Due to contracting issues, such as free agency, player salaries are not an exact measure of player potential. However, they are the only measure of talent which is observable independent of managerial performance for all players. Based on this logic player salaries are used as inputs to the production of winning percentage. This means that the output oriented efficiency will measure how efficiently a manager produced wins given his set of player salaries. This is precisely the measure of performance which is expected to determine whether a manager is retained for the next season.

Due to data limitations and for simplicity of analysis the salaries are divided into offensive salaries and defensive salaries. For offense, the salaries of the players who played the greatest number of games at each infield position and the top three outfielders in games played are summed for each team of each year from 1985 to 2006. Due to the presence of the designated hitter in the American League the analysis must be done separately for the National and American Leagues. When calculating the efficiencies for the American League the designated hitter salary is also included as an offensive salary. For defense, the salaries of the top five pitchers in terms of games started and the top six pitchers in terms of relief appearances are summed for each team of each year. Player salaries are available for the years 1985 on from several online sources, including espn.com, usatoday.com, and the Sean Lahman Baseball Database.

While the potential of a team reflected in its salaries is an important input it is not the only constraint which a manager faces. The level of talent on opposing teams will act as a negative input. If two managers have identical teams and one plays against better competition, that manager cannot be expected to win as much. Therefore, a negative input for the salary of the competition is included in the analysis. Major League teams play the majority of their games against teams within their division. Therefore, the negative input is calculated as the average total salary of the other teams in a given team's division. With this variable included, both measures of the potential a manager is given and the potential he is expected to compete against are taken into consideration.

As any baseball fan knows baseball salaries have grown rapidly over the past several decades. Therefore, in order to compare inputs from different seasons the salaries must be adjusted for the rapid increase in the level of Major League salaries. To accomplish this, a price index is created by calculating the average player salary for each season from 1985 to 2006 and dividing the 2006 average salary by that of the other seasons. This index, presented below, is used to adjust all player salaries into 2006 baseball dollars.

Baseball Salary Price Index	
Year	Multiple
1985	5.9531
1986	6.7943
1987	6.5175
1988	6.2569
1989	5.6078
1990	5.5383
1991	3.1781
1992	2.7059
1993	2.9013
1994	2.7006

1995	2.9374
1996	2.7602
1997	2.3236
1998	2.2130
1999	1.9084
2000	1.4222
2001	1.2433
2002	1.1847
2003	1.1014
2004	1.1376
2005	1.0762
2006	1.0000

With the inputs and outputs defined the output oriented technical efficiency can be calculated. This is done by solving the following maximization problem and taking the inverse of the resulting value of θ. θ can be interpreted as the multiple by which output can be increased using a feasible combination of observed inputs. The subscript 0 identifies values for the team being analyzed. The subscript i identifies the other teams in the comparison group. W, O, D, and C represents winning percentage, offense, defense, and competition respectively.

Maximize θ

Subject to: (1) $\sum \lambda_i W_i \geq \theta W_0$

(2) $\sum \lambda_i O_i \leq O_0$

(3) $\sum \lambda_i D_i \leq D_0$

(4) $\sum \lambda_i C_i \geq C_0$

(5) $\sum \lambda_i = 1$

(6) $\lambda_i \geq 0$

Constraint (1) implies that the combination of other observed winning percentages must be greater than or equal to the observed winning percentage of the manager being evaluated. Constraints (2) and (3) imply that the combination of offensive and defensive inputs must be less than the inputs of the manager under consideration. Equation (4) states that the combination of the negative competition inputs must be at least as great as the competition faced by the manager being evaluated. Constraint (5) implies variable returns to scale by eliminating scaled down versions of one input bundle from the feasibility set. The last constraint simply assures that there are no negative inputs.

The fact that W_0 is a feasible level of output assures that the maximum value of θ will be greater than 1. Therefore, the technical efficiency, or percentage of feasible output which is being produced, is the inverse of θ. Because θ is greater than 1 this number will always lie in the closed interval from 0 to 1. For example, if a manager has an efficiency score of 0.8 his winning percentage is 80% of what could have been achieved by a linear combination of other observations which has the same or lesser inputs.

The output oriented efficiency is calculated for each manager for each season from 1986 to 2005 for both the American and National Leagues. In order to increase the number of comparison input output combinations, for each season each manager is evaluated compared to all managers in that year along with the previous and following seasons. For example, for the Philadelphia Phillies in 2005 Charlie Manuel is evaluated relative to all National League managers from the years 2004, 2005, and 2006. The inverse of the resulting θ from the maximization problem with Charlie Manuel's 2005 season as the evaluated observation is his technical efficiency for that year. These efficiencies are calculated for each manager in the American and National League with more than five consecutive games managed for the years 1986 to 2005 and are used in the subsequent survival analysis. A ranking of managers in terms of efficiency is listed in the following table. The reported technical efficiencies are game weighted averages from the years 1986 through 2005 for all managers with at least 200 games managed within the sample.

Managers 1986-2005 (Minimum 200 Games)							
Rank	Manager	Games in Sample	Average TE	Rank	Manager	Games in Sample	Average TE
1	Ron Gardenhire	647	1.0000	43	John Wathan	646	0.8567
2	John Boles	446	0.9996	44	Jimy Williams	1701	0.8564
3	Bobby Cox	2460	0.9868	45	Tom Kelly	2386	0.8540
4	Ken Macha	486	0.9689	46	Buck Showalter	1554	0.8534
5	Ned Yost	485	0.9666	47	Ray Knight	261	0.8512
6	Jack McKeon	1356	0.9516	48	Jim Lefebvre	859	0.8507
7	Bruce Bochy	1764	0.9462	49	Hal McRae	872	0.8496
8	Larry Dierker	810	0.9390	50	Johnny Oates	1544	0.8490
9	Eric Wedge	486	0.9372	51	Charlie Manuel	573	0.8489
10	Jim Leyland	2202	0.9354	52	Doc Edwards	380	0.8486
11	Felipe Alou	1893	0.9310	53	Joe Morgan	563	0.8474
12	Don Zimmer	524	0.9289	54	Buddy Bell	919	0.8443
13	Bob Brenly	565	0.9283	55	Davey Johnson	1027	0.8418
14	Marcel Lachemann	331	0.9269	56	Cito Gaston	1319	0.8411
15	Art Howe	2266	0.9258	57	Jeff Torborg	994	0.8398
16	Dick Williams	351	0.9199	58	Greg Riddoch	394	0.8366
17	Frank Robinson	1164	0.9166	59	Doug Rader	448	0.8338
18	Pete Rose	236	0.9131	60	Larry Bowa	853	0.8297
19	Mike Scioscia	972	0.9029	61	Lloyd McClendon	782	0.8122
20	Terry Francona	972	0.8985	62	Davey Lopes	340	0.8057
21	Lou Piniella	2939	0.8975	63	Bob Melvin	486	0.8054
22	Gene Lamont	1115	0.8968	64	Rene Lachemann	506	0.8044
23	Larry Rothschild	499	0.8956	65	Carlos Tosca	382	0.8023
24	Dusty Baker	2042	0.8937	66	Tony Muser	748	0.7994
25	Joe Torre	2324	0.8935	67	Terry Bevington	437	0.7914
26	Buck Rodgers	313	0.8934	68	Butch Hobson	439	0.7887
27	Tony LaRussa	3090	0.8920	69	Clint Hurdle	626	0.7845
28	Grady Little	324	0.8869	70	Dave Miley	289	0.7803
29	Jerry Manuel	971	0.8864	71	Buck Martinez	215	0.7780
30	Bill Russell	322	0.8857	72	Nick Leyva	338	0.7746
31	Bud Harrelson	274	0.8844	73	Jim Riggleman	1085	0.7722
32	Kevin Kennedy	582	0.8830	74	John McNamara	267	0.7587
33	Bobby Valentine	1003	0.8826	75	Dallas Green	633	0.7584
34	Tom Trebelhorn	932	0.8713	76	Tony Pena	483	0.7497
35	Hal Lanier	486	0.8687	77	Alan Trammell	486	0.7446
36	Mike Hargrove	2123	0.8671	78	Gene Michael	238	0.7420
37	Don Baylor	1317	0.8666	79	Russ Nixon	347	0.7222
38	Terry Collins	907	0.8622	80	Lee Elia	254	0.7200
39	Phil Garner	1748	0.8616	81	Jerry Narron	389	0.7177
40	Jim Tracy	810	0.8613	82	Ray Miller	324	0.7027
41	Bob Boone	815	0.8605	83	Stump Merrill	275	0.6905
42	Jim Fregosi	1637	0.8585	84	Chuck Tanner	361	0.6622

Survival Analysis

The goal of this paper is to determine whether race has a significant impact on the probability of a manager surviving to the next season. This can be accomplished through the use of survival time analysis. Survival time analysis examines the relationship between the time to an event and several characteristics, referred to as covariates. The goal of survival analysis is to estimate a survival function which gives the probability of survival to a certain time period given a set of covariates. These covariates can be constant over time, such as minority status, or varying each period, such as winning percentage.

Survival time analysis can be done by making distributional assumptions about the survival function. Models which use this method are referred to as parametric models. The first step to estimating such models is to estimate the hazard rate as a function of the covariates. The hazard rate is simply the drop out rate in a given time period conditional on a set of covariates. This rate is always positive so the model is assumed to be linear in the log of the hazard rate. Therefore, the model of interest is the following.

$$\log(h_i) = B_0 + B_1X_{i1} + B_2X_{i2} + \dots + B_nX_{in}$$

As can be seen this model does not depend on time. Therefore, it assumes that the hazard rate is constant over time. This leads to a survival function of the following form.

$$S(t) = e^{-ht}$$

This model is referred to as the exponential survival model and is the most simplistic of the commonly used models due to its assumption of a constant hazard rate over time.

The assumption that the hazard rate is constant over time is often inappropriate and therefore models which allow for the hazard rate to vary over time may be more appropriate. The most common parametric forms of these models are the Weibull and Gompertz models. These models assume that there is some underlying hazard rate which is dependent on time. It is also assumed that there is no interaction between time and the covariates. The covariates effect the hazard rate by proportionally changing the underlying rate for a given time period. This is why these models are referred to as parametric proportional hazards models. Mathematically the hazard rate is modeled as follows.

$$h(t) = h_0(t)\exp(B_0 + B_1X_{i1} + B_2X_{i2} + \ldots + B_nX_{in})$$

Different functional forms of $h_0(t)$ will lead to different survival functions. The underlying hazard rate is commonly assumed to have a Weibull distribution. This leads to the following form for the survival function.

$$S(t) = e^{-(ht)p}$$

This is desirable as the hazard rate will be either increasing or decreasing monotonically with time depending on the value of the estimated parameter p. If p is greater than 1 the hazard rate is increasing over time. If p is less than 1 the hazard rate is decreasing over time. As can be seen if p is equal to 1 this leads to the exponential model where the hazard rate is constant over time.

Another distributional assumption which may be appropriate is the Gompertz distribution. If the underlying hazard is assumed to follow a Gompertz distribution the resulting survival function is of the following form.

$$S(t) = \exp[(h/r)(1-e^{rt})]$$

Under this assumption the hazard rate will either increase or decrease at an exponential rate depending on the value of the estimated parameter r. These three distributional assumptions can be used to estimate a hazard rate which is constant over time, increasing over time, and increasing exponentially over time. These three models are estimated using maximum likelihood and the results are presented in the next section.

An alternative approach to these parametric models is to estimate a Cox proportional hazards model. In order to estimate this model the ratio of hazards for two observations is taken as follows.

$$h_i(t)/h_j(t) = h_0(t)\exp(B_0 + B_1X_{i1} + \ldots + B_nX_{in}) / h_0(t)\exp(B_0 + B_1X_{j1} + \ldots + B_nX_{jn})$$

$$h_i(t)/h_j(t) = \exp(B_0 + B_1X_{i1} + \ldots + B_nX_{in}) / \exp(B_0 + B_1X_{j1} + \ldots + B_nX_{jn})$$

Due to the fact that the baseline hazards are independent of the covariates the baseline hazards cancel leaving a hazard ratio which is independent of time. Despite the fact that the underlying hazard function is not defined the model can still be estimated by the method of partial likelihood. This method is presented by Cox in the 1972 paper in which he first introduces the Cox model. While these models are not as efficient as a correctly specified parametric model they do not depend on distributional assumptions. This avoids the risk of obtaining misleading results due to an incorrectly specified parametric model. A Cox proportional hazards model is estimated in addition to the three parametric models and the results are presented in the following section.

Estimation Results

The data on Major League Baseball managers from 1986 to 2005 consist of 573 observations for which the individuals managed more than 5 consecutive games. These observations are used to construct managerial streaks which range in duration from less than one season to 16 seasons. In order to avoid any left censoring issues only streaks which began from 1986 on are included in the analysis. This results in a sample of 180 managerial streaks. Minority managers account for 24 of these streaks which are distributed between 16 different teams. Using the performance and personal characteristic variables as covariates the four survival time models previously discussed are estimated. These models are estimated for numerous different combinations of covariates in order to determine which variables have a statistically significant effect on survival time. The estimation results with all variables included are presented below.

Estimation Results With All Variables Included

	Exponential		Weibull		Gompertz		Cox Prop. Hazards	
	Hazard Ratio	P-Value	Hazard Ratio	P-Value	Hazard Ratio	P-Value	Hazard Ratio	P-Value
Minority	0.7069960	0.166	0.5330281	0.014	0.5996457	0.046	0.6385576	0.078
Efficiency	0.0795495	0.007	0.0536909	0.002	0.0468547	0.002	0.0731800	0.006
Winning %	0.9998889	0.945	1.000437	0.791	1.000668	0.687	1.000088	0.957
Playoff Wins	0.9629778	0.404	0.9597056	0.368	0.9678998	0.472	0.9622753	0.399
Experience	0.9999485	0.877	0.9985763	0.000	0.9994366	0.104	0.9995162	0.192
Experience^2	0.9999999	0.499	1.000000	0.115	0.9999999	0.569	1.000000	0.854
Age	1.033812	0.788	0.9786913	0.867	1.000140	0.999	1.035281	0.780
Age^2	1.000021	0.986	1.000730	0.549	1.000473	0.694	1.000043	0.971
p	-		1.898794		-		-	
r	-		-		0.18272		-	

Under all but one model specification the coefficients on age squared and experience squared are statistically insignificant at any conventional level. Additionally, for most model specifications the squared terms have the same sign as experience and age. Therefore, it is concluded that age and experience do not have diminishing effects and the squared terms are dropped from the final model.

Winning percentage does have the anticipated positive effect on survival under all specifications. However, this effect becomes highly insignificant when technical efficiency is added to the model. Without technical efficiency the coefficient on winning percentage has a p-value of less than .001. When technical efficiency is added to the model, this p-value jumps to over .6 for all of the model specifications. This implies that winning percentage does not provide any additional information to the model when technical efficiency is included. .

Similarly, playoff wins have the anticipated positive effect on survival. However, for most specifications this effect is not significant at conventional levels. As with winning percentage, for all specifications playoff wins become highly insignificant when technical efficiency is included in the model.

Winning percentage and playoff wins are both correlated with managerial efficiency with correlation coefficients of .75 and .27 respectively. This correlation may be contributing to their insignificance when efficiency is included in the model. However, the inclusion or exclusion of these variables does not significantly change the magnitude or significance of the coefficient on the variable of interest, minority status. Therefore, winning percentage and post season wins are excluded from the final model chosen.

The exponential, Weibell, Gompertz, and Cox proportional hazards models are estimated with technical efficiency, age, experience, and minority status as covariates. The resulting hazard ratios and p-values are presented below. A hazard ratio greater than 1 implies that that covariate has a positive impact on the baseline hazard rate. A hazard ratio of less than 1 implies that that covariate has a negative impact on the baseline hazard rate.

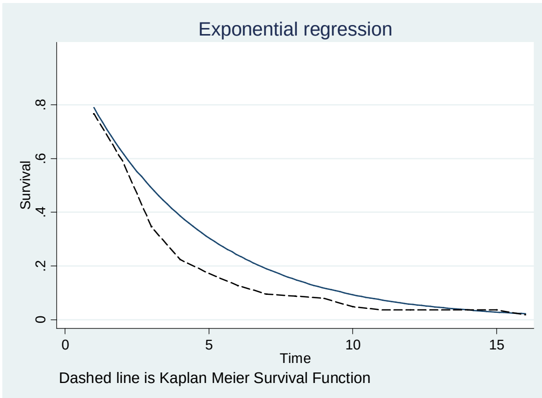
Estimation Results for the Selected Model

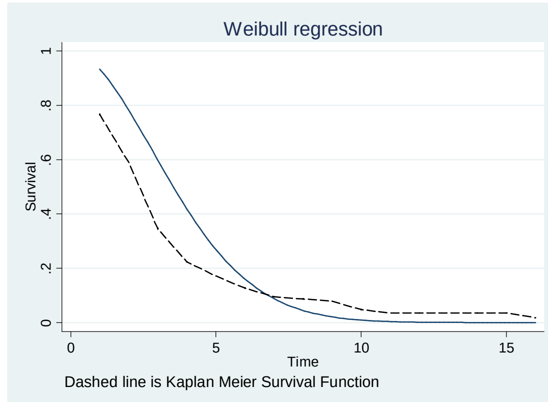
	Exponential		Weibull		Gompertz		Cox Prop. Hazards	
	Hazard	P-Value	Hazard	P-Value	Hazard	P-Value	Hazard	P-Value

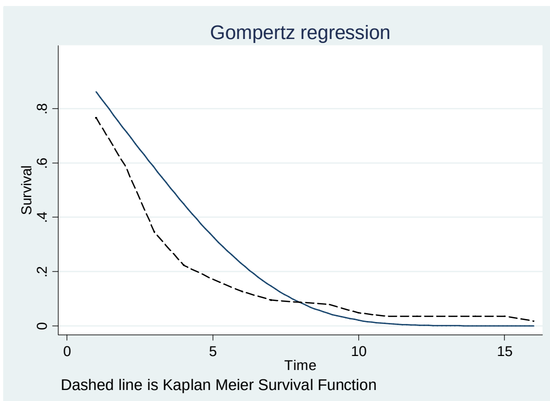
	Ratio		Ratio		Ratio		Ratio	
Minority	0.71391	0.176	0.52267	0.011	0.60326	0.047	0.63369	0.072
Efficiency	0.07379	0.000	0.06400	0.000	0.06786	0.000	0.07293	0.000
Experience	0.99972	0.031	0.99908	0.000	0.99926	0.000	0.99956	0.006
Age	1.03797	0.007	1.05566	0.000	1.05202	0.000	1.04072	0.005
p	-		1.830658		-		-	
r	-		-		0.181839		-	

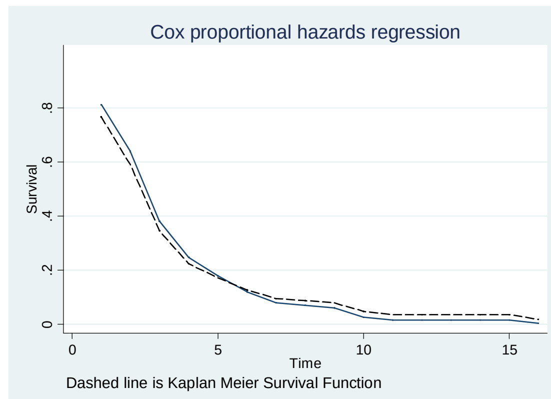
As can be seen in the regression results all of the covariates are statistically significant at the 5 percent level with the exception of minority status for the exponential and Cox models where it is significant at the 18 and 8 percent levels respectively. The results are also similar in sign and magnitude for all four models. Minority status, efficiency, and experience all appear to increase the probability of survival while age has a negative effect on the probability of survival.

In order to evaluate which model is most appropriate the survival functions for all four models are plotted along with the Kaplan-Meier survival function. The Kaplan-Meier survival function is a description of the observed survival rates which treats observations for which no failure is ever observed as having survived in the last period they are observed. However, it does not include these observations in the number of observations which are subject to failure in the next period. The survival probabilities for each period are then used to calculate the probability of survival past a given period as a compound conditional probability. This survival function can be interpreted as the observed survival function against which models should be compared.









Clearly the Cox proportional hazards model does the best job of approximating the Kaplan Meier survival function. Under this model all covariates are significant at the 1 percent level except for minority status which is significant at the 8 percent level.

Specification Tests

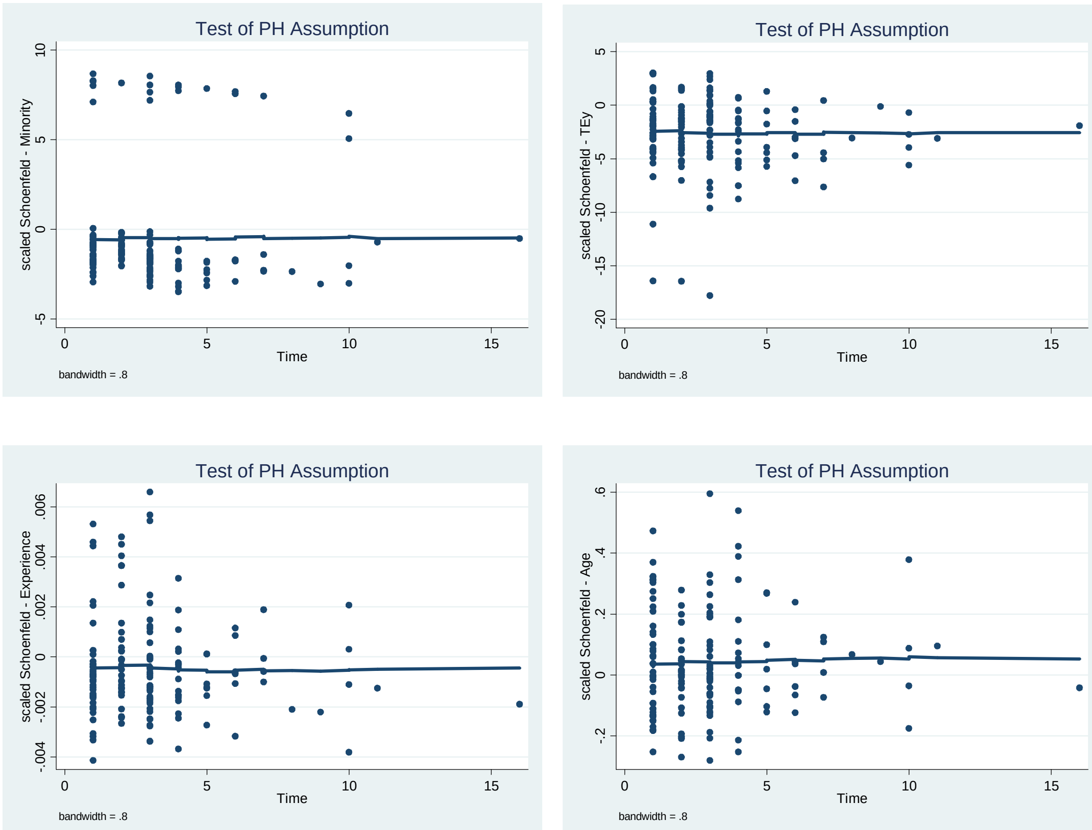
While the advantage of the Cox model is that it does not rely on distributional assumptions about the survival function, the assumption of proportional hazards must be tested before the Cox model is accepted as appropriate. One test of this assumption is to estimate the model with interaction terms of time and each covariate included in the model. If any of these interaction terms are statistically significant it is evidence that they violate the assumption of proportional hazards.

	Cox Prop. Hazards	
	Hazard Ratio	P-Value

Minority	0.49310	0.169
Efficiency	0.13248	0.008
Experience	0.99971	0.297
Age	1.04297	0.090
t*Minority	1.06696	0.634
t*Efficiency	0.78124	0.362
t*Experience	0.99996	0.573
t*Age	0.99864	0.853

The p-values on all four interaction terms are greater than .3. This does not provide any evidence against the assumption of proportional hazards. Another test of whether the assumption of proportional hazards is violated is to run a generalized linear regression of the scaled Schoenfeld residuals on time. As with the previous test a significant coefficient on time for any of the covariates is evidence that the proportional hazards assumption is violated. Graphically, this is equivalent to having slopes equal to zero in the graphs of these regressions.

Test of proportional hazards assumption		
	rho	P-Value
Minority	0.04675	0.562
Efficiency	-0.05657	0.591
Experience	-0.08468	0.290
Age	0.00909	0.914
Global Test	-	0.677



The coefficient on time is highly insignificant for all covariates individually and jointly. This result is consistent with the lack of slope in the graphs of the residuals versus time. Therefore, there is no evidence that the proportionality assumption is violated and the Cox proportional hazards model is appropriate.

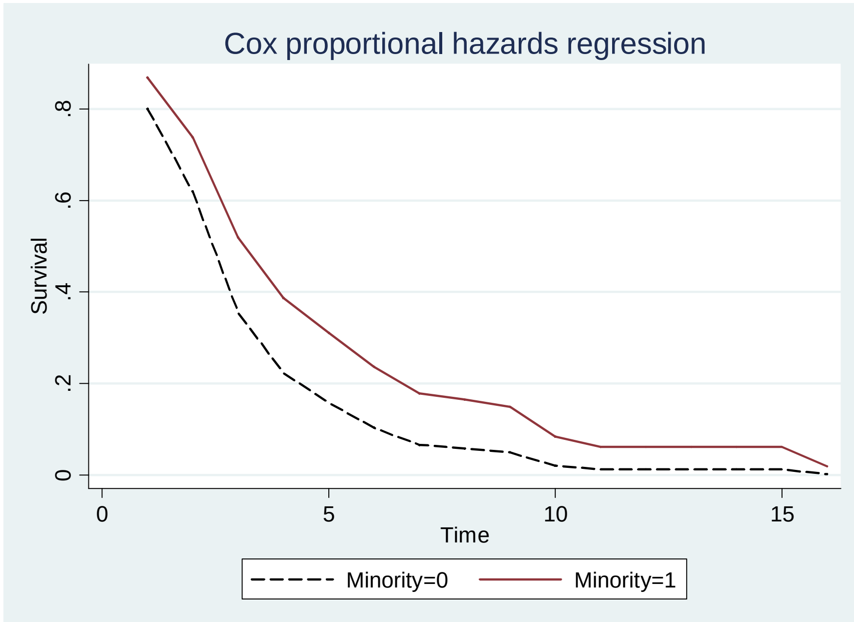
Given the observed statistically significant relationship between minority status and survival, it is also possible that the covariates affect minorities differently than white managers. In order to test this hypothesis, interaction terms of minority status with efficiency, experience, and age are added to the models previously estimated. It is found that the coefficients on these interaction terms are insignificant at conventional levels for all model specifications. Therefore, it is concluded that the covariates do not affect minorities differently and the previously estimated model is appropriate.

Conclusions

The model selected is a Cox proportional hazards model with minority status, efficiency, experience, and age as covariates. In this model managerial efficiency increases the likelihood of survival as expected and is significant at a confidence level of less than 1 percent. It is interesting to note that winning percentage and playoff wins are highly insignificant when added to the model. This implies that it is not important how many games a manager wins but rather how well he converts his given resources into wins. This makes sense as owners of teams are expected to maximize profits not wins. Therefore, they will retain a manager with low wins if that manager did the best possible with his given resources.

Experience also increases the probability of a manager returning the following season. This is to be expected as experience managing will add knowledge to a manager increasing his human capital. The more human capital a manager possesses the less likely it is that someone more qualified can be found to replace him. Unlike experience, age decreases the probability that a manager will return the next season. This is not surprising as older managers are more likely to retire. It is also likely that older managers will be unable to relate to players who are significantly younger. The added experience from age is also likely captured by the experience term resulting in an overall negative relationship between age and the probability of returning.

The chosen model estimates a positive relationship between minority status and survival. This impact is significant at the 8 percent level. To see the impact of minority status on managerial retention the estimated survival function is plotted with minority equal to zero and one. The other covariates are evaluated at their sample mean.



Marginal Effect of Being a Minority on the Probability of Retention	
Year of Tenure	Marginal Effect
1	0.068
2	0.076
3	0.130
4	0.117
5	0.096
6	0.111
7	0.114
8	0.042
9	0.053
10	0.159

As can be seen in the preceding graph and table, being a minority increases the probability of returning the next season. On average during the first ten seasons of tenure minority status increases the probability of retention by 9.6 percentage points. This effect is substantial and statistically significant with a p-value of .072. The positive relationship between minority status and survival also appears to be robust to different model specifications and combinations of variables. This result leads to the conclusion that minority status does not decrease the probability of a manager returning. Rather, there is significant evidence that minority status actually increases the probability of retention. Therefore, the current decrease in the number of minority managers is not due to discrimination in the retention decision.

While the analysis presented here finds no discrimination in the decision to keep a current manager, the results may be motivated by discrimination in hiring. If there is discrimination against minorities in the hiring of major league managers, then only the most highly qualified minorities will become managers. Therefore, the minority managers who have the possibility of not returning may possess some unobserved characteristic which enabled them to overcome the discrimination in hiring. It is likely that this unobserved characteristic may also make the manager less likely to be fired and thus cause the relationship observed in this analysis. However, this analysis attempts to account for all of the managers' characteristics by controlling for performance, experience, and age. After justifying the included variables and the appropriateness of the model this analysis concludes that the relatively low number of minority managers in the major leagues is not due to a lower probability of survival.

Works Cited

Anderson Torben and Sumner J. La Croix. 1991. “Customer Racial Discrimination in Major League Baseball.” *Economic Inquiry* 29: 665-677.

Audas, Rick; Stephen Dobson; John Goddard “Organizational Performance and Managerial Turnover” *Managerial and Decision Economics*, Vol. 20, No. 6. (Sep., 1999), pp. 305-318.

Banker, R. D., Charnes, A., and W. W. Cooper. (1984). “Some models for estimating technical and scale inefficiencies.” *Management Science* 39, 1261-1264.

Charnes, A., Cooper,W., and Rhodes, R. (1978). “Measuring the efficiency of decision making units.” *European Journal of Operational Research*, *2*, 429-444.

Cox, D. R. 1972. “Regression Models and Life Tables (with Discussion).” Journal of the Royal Statistical Society, Series B 34:187—220.

Cox, D. R. & D. Oakes. 1984. “Analysis of Survival Data.” London: Chapman and Hall.

Einolf, Karl. (2004). “Is Winning Everything?: A Data Envelopment Analysis of Major League Baseball and the National Football League.” *Journal of Sports Economics*, *5*, 127-151.

Evans, Arthur S. Jr. “Blacks as Key Functionaries: A study of Racial Stratification in Professional Sport” *Journal of Black Studies*, Vol. 28, No. 1. (Sep., 1997), pp. 43-59.

Fizel, J. L., & D’Itri, M. (1996). “Estimating managerial efficiency: The case of college basketball coaches.” *Journal of Sport Management*, *10*, 435-445.

Fizel, John L.; Michael P. D’Itri. “Managerial Efficiency, Managerial Succession and Organizational Performance.” *Managerial and Decision Economics*, Vol. 18, No. 4. (Jun., 1997), pp 295-308.

Hanssen, Andrew and Torben Andersen. “Has Discrimination Lessened Over Time? A Test Using Baseball’s All-star Vote.” *Economic Inquiry*, **37**, (1999)pp. 326-352.

Jewell, R. Todd, Robert W. Brown, and Scott E. Miles. “Measuring Discrimination in Major League Baseball: Evidence from the Baseball Hall of Fame.” *Applied Economics*, January 2002, 34(1), pp.167-177.

Jiobu, Robert M. “Racial Inequality in a Public Arena: The Case of Professional Baseball.” *Social Forces*, Vol. 67, No. 2. (Dec., 1988), pp. 524-534.

Kahn, Lawrence M. “Managerial Quality, Team Success, and Individual Player Performance in Major League Baseball.” *Industrial and Labor Relations Review*, Vol. 46, No. 3. (Apr., 1993), pp. 531-547.

Lin, D. Y. and L. J. Wei. “The robust inference of the Cox proportional hazards model.” *Journal of the American Statistical Association* 84: 1074–1078.

Lapchick, Richard, Boma Ekiyor and Horacio Ruiz. “The 2006 Racial and Gender Report Card: Major League Baseball” The University of Central Florida’s Institute for Diversity and Ethics in Sport (2006).

Porter, Philip K.; Gerald W. Scully. “Measuring Managerial Efficiency: The Case of Baseball.” *Southern Economic Journal*, Vol. 48, No. 3. (Jan., 1982), pp. 642-650.

Scully, Gerald W. “Managerial Efficiency and Survivability in Professional Team Sports.” *Managerial and Decision Econoimcs*, Vol. 15, No. 5, Special Issue: The Economics of Sports Enterprises. (Sep.-Oct., 1994), pp. 403-411.

Therneau, T. M. & P. M. Grambsch. 2000. “Modeling Survival Data: Extending the Cox Model.” New York: Springer.