Does mid-season change of coach improve team performance? Evidence from the history of the NBA

Date of submission: 20 December 2010

Jose A. Martínez

Departamento de Economía de la Empresa

Universidad Politécnica de Cartagena

Email: josean.martinez@upct.es

Este es un Working Paper. Por favor, no citar este documento sin el permiso del autor.

Cualquier sugerencia o comentario es bienvenido.

Does mid-season change of coach improve team performance? Evidence from the history of the NBA.

Abstract

This research has advanced in the understanding of the effect of hiring new coaches on performance of NBA teams, when change is achieved in the middle of the season. Changing a coach is an important managerial decision which does not guarantee improvement of results in the short term, i.e. in the same season. Using an easily understandable procedure based on comparison of proportions in a finite population approach framework, we show that only about 15% of new coaches outperformed in a significantly way their predecessors. In order to maximize the probability of success, highly experienced coaches, with a long career as former NBA players should be signed. In addition, change should be made before season advances. These three factors slightly contribute to increase the probability of success, being the most important variable the winning percentage of the team at the moment of change. Therefore, worse teams are more probable to be successful than better teams when a change is made. Finally, implications and limitations are discussed.

Keywords: Changing a coach, team performance, basketball, decision making in sports

Does mid-season change of coach improve team performance? Evidence from the history of the NBA

One of the most important management dilemmas for owners and general managers of sports teams is about changing a coach when results are not as good as expected within a season. Numerous studies have addressed this topic in football/soccer (e.g. Barros, Frick & Passos, 2009; Bruinshoofd & ter Weel, 2003; Frick, Pestana & Prinz, 2010; Koning, 2003; González-Gómez, Picazo-Tadeo & García-Rubio, 2011; Salomo, Teichmann & Albrechts, 2000; Tena & Forrest, 2007; Van Dalen, 1994). In other sports, this topic has been also a matter of subject, but has attracted lesser attention: American football (Brown, 1982; McTeer, White & Persad, 1995), baseball (Gamson & Scotch, 1964; McTeer, White & Persad, 1995; Scully, 1995), hockey (McTeer, White & Persad, 1995) or basketball (Fizel & D'Itri, 1999; Giambatista, 2004; McTeer, White & Persad, 1995; Scully, 1995). Results of these studies are contradictory (see Koning, 2003, and González-Gómez, Picazo-Tadeo & García-Rubio, 2011). One of the reasons of these disparate findings is the heterogeneous nature of data (disparate sports, seasons, competitions…), and the distinct assumptions, models and statistical tools used to analyse such data.

Changing a coach is an important managerial decision because it has several consequences for teams: (1) short-term financial costs: owners have to compensate the coach who has been dismissed and to pay to the new hired coach[1]; (2) team performance uncertainty: there are inconsistent results about if changing a coach

[1] For example, Pistons paid about $6 million dollars to Larry Brown in 2005 after firing him (http://www.highbeam.com/doc/1P2-13896726.html). In 2008, Kings paid three head coaching salaries that season to Natt, Theus and Musselman (http://nbcsports.msnbc.com/id/28239643)

improves, declines or maintains the current performance; then it is highly complicated to make a reliable prediction about the success of the potential change.

Regarding the first point, salaries of NBA coaches are the highest for North American professional sports (Van Ripper, 2010). The average NBA coaching salary grew from $2.73 million in 1998 to $3.81 million in 2007[2], being about $3.4 million in 2010 (García, 2010). It is true that the average player salary is $5.9 million, meaning the majority of coaches make less than those they're paid to lead (García, 2010), but this does not mean that these quantities paid to coaches are without importance for team budgets.

Regarding the second point, as Frick, Pestana and Prinz (2010) explains[3], some theories have been used to analyze either CEO or head coach turnover: ‘‘common sense theory”, ‘‘vicious circle theory” and, ‘‘ritual scapegoating theory” Common sense theory is based on the assumption that a new CEO or a new head coach will be hired if he has the required expertise and experience to increase the performance of the firm or the team. Therefore, a new head coach can be instrumental in breaking organizational inertia and then in initiating strategic change. Vicious circle theory assumes the opposite; consequently, successions are likely to have a disruptive effect on the team in terms of increasing instability and ambiguity and will thus lead to an even poorer performance. At last, ritual scapegoating theory suggests that there is no relationship between succession and performance and that succession events just serve as signals to stakeholders that required organizational change is under way.

---

[2] See http://www.gastongazette.com/sports/million-3839-salary-year.html
[3] See also Fizel and D’itri (1999)

Therefore, both theory and practice do not provide a clear answer to the question: changing a coach in the middle of a season improves team performance? In addition, in professional basketball, there is a lack of studies regarding this topic. As stated previously, the studies of McTeer, White & Persad (1995), Scully (1995), Fizel and D'Itri (1999) and Giambatista (2004) are the most relevant. However, McTeer, White & Persad (1995) imposed several restrictions to the characteristics of changes to be included in the sample, obtaining a very small number of cases (22) to analyse. With regard to Scully (1995), he focused his research about changes of coaches among consecutive seasons, and not within a season. On the other hand, in the research of Fizel and D'itri (1999) most of the changes were end of season with relatively few mid-season changes. When a mid-season change occurred, the performance of the fired coach was then compared to his permanent not temporary successor. Regarding Giambatista's (2004) study, he mainly focused on the analysis of life-cycle of coaches in the NBA, using all the changes of coaches occurred in the NBA till the 2001-02 season. Results of this study indicated that midseason hires were associated with a sharp performance decline; team performance following an off-season hire was higher than performance following a midseason hire. Therefore, Giambatista (2004) did not explicitly compare the performance of hired coaches against fired coaches within the same season.

In this research, we analyse all mid-season coaches changes in the history of the NBA, from the first change (1949-50 season), to the latter (2009-10 season). We statistically compare the winning percentage of each team when a coach is replaced[4], with the winning percentage got by the same teams managed by new coaches, in order

[4] We consider all changes of head coaches, regardless of these changes were voluntary (resigned coaches) personal or not (fired coaches).

to obtain a reliable test considering performance achieved in a sample of the whole games of a season. In addition we test a model in order to explain the successes cases, i.e. when a new coach got statistically better results than the coach he replaced. To our knowledge, this is one of the few studies in sports that reviews all the mid-season changes of coaches in a professional league, because other studies have focused only in a specific period of time. For example: Van Dalen (1995), 1 season; Tena and Forrest (2007), 3 seasons; Koning (2003), 4 seasons; Fizel and D'Itri (1999), 8 seasons; González-Gómez, Picazo-Tadeo and García-Rubio (2011), 8 seasons; Salomo, Teichmann and Albrechts (2000), 19 seasons; Frick, Pestana and Prinz (2010), 22 seasons. We review all of the 61 seasons of the NBA league. It is true that Horowitz (1994) made a similar extensive review in baseball (from 1903 to 1992), but he aggregated both between and within season changes, and only provided results of performance comparisons for a random sample of 25 teams. In addition, Giambatista (2004) and McTeer, White & Persad (1995) made also an akin review for the NBA league, from the beginning of the competition to the 2001-02 season and 1988-89 season, respectively. However, the former study mainly focused of life cycle of coaches, and the second study included a very small number of cases to analyse.

Results from our research show that changing a coach only improves performance in about 15% of cases, being the neutral effect prevalent in the majority of cases (about 75%). In addition, the probability of success (i.e. the probability of a new coach improves performance), increases with his experience and expertise (defined by the number of games managed and the number of wins got till the moment of change) and also with his experience as a former NBA player (defined by the number of season played), but it inversely related to the winning percentage of a team at the moment of change, and the difference between games played by old and new coaches.

Therefore, the unique contribution of this research to sport management literature is threefold: Firstly, we propose a simple form to compare performance of new coaches with performance of replaced coaches, based on principles of basic probability, in order to study the success of decisions made by owners and general managers. This method may be easily understood and implemented for sport managers; Secondly, we make the most extensive review of changes of coaches in basketball, analysing all the mid-season changes occurred in the history of the NBA, providing a full description of these changes. And thirdly, we identify factors influencing the probability of success, providing practical implications for decision makers about the effect of such variables in the change of the probability of success.

Method

Data and measures

We collected data from www.basketball-reference.com, the major source of basketball data available at the present time. We considered all mid-season changes of coaches from the first NBA season (1949-50) to the last (2009-10). We also registered other measures: (1) experience and expertise of new coaches (winning percentage and games managed in the league at that moment, and winning percentage and games managed in the league in their whole career). In addition we considered data for coaches who were former NBA players (number of seasons and minutes played, average winshares[5] per 48 minutes, if they were all-star players and if they are Hall of Fame players); (2) number of games a team played at the moment of change, at the end of the season, and played after new coaches were signed; (3) winning percentage of teams at the moment of change, at the end of the season, and obtained by new coaches.

Procedure to compare performance

We compared the winning percentage got by teams trained by old coaches with the winning percentage obtained by teams trained by new coaches. We used the finite population approach of the binomial distribution for comparing two proportions (see Levy and Lemeshow, 1999). As the winning percentage is a variable ranged from 0 to 1, we considered the value of such variable in a specific moment of the season as p, which is an estimate of P, i.e. the winning percentage of the team at the end of the season, after playing N games. Recall that the value of N depends of the season, and ranges from 64 of the 1950/51 season to 82 at the current time.

Taking basic principles of statistical inference, to the extent that the n games played by a NBA franchise when a coach left the team approximates N, then the imprecision of p decreases, but p will always has an associated confidence interval when $n < N$ and $0 < p < 1$ Therefore, if we denote p_1 the winning percentage obtained by a coach in n_1 games, and p_2 the winning percentage obtained by his substitute in n_2 games, then we may make a test for the difference of proportions. Therefore, under the null hypothesis of equally of population parameters $P_1 = P_2$, we may compute an approximate 95% confidence interval for the difference of proportions:

[5] Winshares is a measure of productivity achieved by players, which considers offensive and defensive performance, and it is an estimate of the number of wins contributed by a player (http://www.basketball-reference.com/about/ws.html)

$p_1 - p_2 \pm 1.96\sqrt{(S_1^2 + S_2^2)}$, being S_1^2 and S_2^2 the sample variances. These variances are computed using the finite population factor, in the following way for S_1^2:

$\frac{N - n_1}{N} \frac{p_1(1 - p_1)}{n_1 - 1}$, and similarly for S_2^2.

Therefore, if a computed confidence interval contains zero, then we can not reject the hypothesis of equality of proportions, i.e. we can not consider that both winning percentages differ. In addition, the positive or negative value of both extremes of a confidence interval favours the alternative hypothesis of difference in proportions, which means that a replaced coach got significantly better results, or a new coach obtained significantly better performance, respectively.

This simplistic method has the advantage of consider uncertainty in the value of the winning percentage of teams. When a coach only manages a portion of all of the possible games of a season, it is highly desirable to depict the winning percentage obtained using an interval estimation, because we are not sure about if this punctual value at the moment of change reflects without error the strength of the team. This fact has not been addressed in past research on basketball, such as Fidel and D'itri (1999), Giambatista (2004) or McTeer, White and Persad (1995). However, we consider this issue as crucial for making comparisons among performance of different coaches when analysing mid-season changes, because we are comparing two samples of games from two hypothetical populations. Therefore, a statistical test is necessary to ascertain weather both populations are the same or not, i.e. if both winning percentages differs or are statistically the same. In addition, as the criteria to compare winning percentages is based on a probabilistic perspective, we believe that outperforms the procedure achieved by McTeer, White and Persad (1995), authors who subjectively considered 11 games played as a cut-off value for including cases in their study.

Statistical model

Once obtained which coaches improved team performance we were interested in explaining such success. We created a new dichotomous variable y_i for all the i cases, being 0 "no success" (including the cases of no difference in performance and better performance of old coaches), and 1 "success" (better performance of new coaches). In addition, in order to explain such variation in performance, we identified variables which could influence the probability of success. Following the principles of General Linear Modelling, we built a logit model with the following specification:

$\ln\left(\frac{y_i}{1-y_i}\right)=\beta_0+\sum_k \beta_k X_{ki}+e_i$, where X_{ki} was a set of predictors, β_0 and β_k were the coefficients to be estimated, and e was a random error with zero mean which was uncorrelated with predictors.

We considered the following predictors: (1) the experience/expertise[6] of new coaches, reflected in games played in the NBA and the number of wins obtained, X_1; (2) the experience/ expertise of new coaches as former NBA players X_2; (3) the winning percentage of teams at the time of change, X_3; (4) the difference between the percentage of games played in a season between a new and an old coach, X_4; (5) the NBA trajectory of new coaches in the next years, X_5. This latter variable should not be included in predictive models because, obviously, it is not available at the moment of

[6] We speak about experience, expertise or ability of coaches in a similar way. We understand that may be considered different concepts for some authors (e.g. Giambatista, 2004), but we only want to reflect the curriculum vitae of coaches defined by games managed and wins obtained, i.e. the a priori potential or quality of hired coaches.

change, but it must be included in the exploration of explaining models, because it might improve the explained variance of the dependent variable.

Regarding X_1, Salomo, Teichmann and Albrechts (2000), and Giambatista (2004) considered this variable as a good proxy for valuating coaches abilities. In addition, Barros, Frick and Passos (2009) found that the probability of being dismissed was negatively affected by head coaches' experience and their winning percentage. We consider a novel approach to deal with this variable. We used a sigmoid function $\frac{1}{1+e^{-\lambda A_1}}$ (Schmueli, Patel & Bruce, 2007) in order to transform the distribution of the number of games played by a coach into a S-shaped curve. In this case A_1 refers to the number of NBA games played by a coach till the moment of signing with the new team, and λ is a parameter to adjust the curve to the desired rank of values. We hypothesised that this form of curve reflects the effect of experience on performance, Therefore, increasing the number of games played yield a small effect of performance in the first steps of a coach career and when a coach has played a great amount of games.

In order to create an easily interpretable index after the sigmoid transformation, we made the following: First, we considered the distribution of the number of games played by all the 203 coaches. This was a highly asymmetrical distribution with a high amount of zeros (new coaches without any previous experience in the NBA). Taking the form of this distribution, we computed its median (209 games), and then the deviation of each value from that median. These values ranged from -209 (rookie coaches) to 2197 of the Hall of Famer Lenny Wilkens. Second, we calibrated the sigmoid function using the λ parameter, in order to obtain a normalized [0,1] function. This calibration process yielded a $\lambda = 1/75$. Figure 1 shows the final curve. We strongly believed that this S-shaped curve was a better model for characterizing coaching experience, instead of other forms of learning gains, such as the natural logarithm transformation achieved by Giambatista (2004). Logarithm yields a curve with increased diminishing returns to the extent that experience grows. However, we think that in the first steps of NBA coaches career, gains of learning should be modelled using this type of exponential growth.

--- Figure 1 about here ---

We achieved exactly the same procedure with the number of wins got by coaches, because obviously it is not the same to play many games and lose the majority of them than to play many games winning a great percentage of them. Giambatista (2004) also used wins got by coaches, although with the natural logarithm transformation. We hypothesized that "winners" coaches should have a greater experience index (i.e. potential) than "losers". Median of the distribution was 84.5, and deviation from median ranged from -84.5 (rookie coaches) to 1207 of Lenny Wilkens. In this case, the calibration process yielded a $\lambda = 1/20$. Considering that correlation between both measures of coaches experience (once transformed in a normalized form) was about 0.98, and in order to avoid multicollineality problems, we aggregated both measures to finally get an index of coach experience ranged in a [0,2] interval.

An important decision regarding the nature of this variable was about considering experience in the BAA (Basketball Association of America) and ABA (American Basketball Association) leagues. BAA league was a professional basketball league founded in 1946 and, together with ABL (American Basketball League) and

NBL (National Basketball League), contributed to the expansion and promotion of basketball before NBA was created. Moreover, NBA was created from a merge of BAA and ABL. There are only BAA data available (from 1946/47 to 1948/49 seasons), so we decided to include the experience of coaches who trained BAA teams in the same form as if they had trained NBA teams. However, we do not have data regarding experience of coaches in the other professional leagues, such as NBL and ABL.

A similar problem appeared with the ABA league, a professional competition founded in 1967. This league challenged the hegemony of NBA during several years, and finally merged with NBA in 1976. Four ABA teams joined to NBA, so we also decided to include the experience of coaches who trained ABA teams in the same way of NBA teams.

Regarding X_2, Goodall, Kahn and Oswald (2010) found that former star players make the best coaches. In addition, this expert knowledge effect is large. Therefore, we hypothesized that the experience of coaches as former players increased the probability of success. Goodall, Kahn and Oswald (2010) used several measures of experience. Following these authors, we also used measures such as the number of season played, if the player was All-Star, and if the player is member of the Hall of Fame. In addition, and as a novelty, we registered information regarding number of minutes played and winshares, because we thought that these variable could measure the quality of experience of players in a more reliable way. Note that players could play many NBA seasons but with a marginal presence in the roster rotation, so minutes played and a productivity index such as winshares could overcome this limitation.

We, thus, created three indicators for measuring X_2. The first one was an ordered variable ranged from 0 to 3, being the codification as follows: "0" coaches without NBA experience as players; "1", coaches with NBA experience; "2" coaches who had been All-Star as players; and "3", coaches pertaining to the Hall of Fame as players. The second variable was the number of seasons played as players. And the third variable was a combination of number of minutes played and the winshares per 48 minutes achieved. As we thought S-shaped curve was the best way to characterize ability and performance variables, then we transformed the second and third variable using the procedure previously depicted for X_1. However, we also tested the alternative specification of consider the first two variables in the original scales, as Goodall, Kahn and Oswald (2010) did. Therefore, we ran several model using these three indicators, in order to choose the indicator which the best explanatory power.

Regarding X_3, we hypothesized that the winning percentage of a team at the time of change would be negatively associated with coach performance improvement, i.e. new coach success. The rationale of this simple reasoning is straightforward: it will be easier to outperform a very bad winning percentage than a good winning percentage. Obviously, this association would be moderated for the intrinsic quality of each team considered, because it would be easier for coaches of good teams to outperform a very bad winning percentage than for coaches of bad teams. However, as the proxy variable for valuating quality of teams is the winning percentage at the end of the season, we did not consider a good decision to use it as predictor, because the dependent variable is precisely measuring change in quality of teams due to the new coach effect.

Regarding X_4, the difference between the percentage of games played in a season between a new and a replaced coach, this variable should be included in order to study if the probability of success increases with the parity of games played by replaced and new coaches. To the extent that n_1 approximates n_2 power of the test would increase, considering constant the remaining factors, so it would be more probable to obtain significant results. In addition, this variable could provide information regarding a “time effect”, i.e. if managing more games improves performance.

Finally, regarding X_5, the NBA trajectory of new coaches in their whole career could serve as an additional information to calibrate the quality of coaches. A coach with a prominent NBA career after being signed by a team, could be indicative of the quality of this coach at the moment of change. It is certain that this reasoning may be criticized, because a coach could improve his potential to the extent that he increases his experience, so we will only explore the behaviour of this variable in the model. We built this variable using exactly the same procedure of X_1, being anchored in a [0,2] interval. Note that correlation between X_1 and X_5 was 0.55. Although this can be considered a high correlation, it is also true that indicates that there is an important portion of uncommon variance between both variables.

Omitted variables

With regard to model specification, two important issues were not addressed: the home/away calendar and the strength of schedule. Both elements have been considered important to study the effect of firing a coach on team performance (Koning, 2003), because of the effect of home advantage (see Winston, 2009) and the effect of disparate winning percentages on the probability of win (see Huang, Weng & Lin, 2006).

However, there are important difficulties to consider these factors in our study. The main limitation arises from the practical impossibility to collect data about the winning percentage of each team playing against all of the cases we have considered during 61 years of competition. Note that it would be necessary to "manually" register the winning percentage of rivals and if the game was played at home or away for the approximately 6842 games played by teams before changing a coach, and the approximately 7939 games played by the same teams after signing a new coach, i.e. a total of 14781 games[7]. In addition, there is interaction between the home/away variable and the winning percentage variable, so it had not been feasible to aggregate the winning percentage of teams played against the focused team that changed a coach in a season (this fact would have facilitated the task), because the home/away advantage acts in a different form for teams having better performance. As Koning (2003) claimed, the effect of home advantage variable varies between teams. We made a small analysis in order to ascertain this in basketball. We got data from three seasons (from 2006/07 to 2008/09), and computed the number of total wins for each team and the percentage of home wins. The mean percentage of home wins was 0.61, indicating the influence of this factor for winning a game (11 points above the neutral value of 0.5). However, the distribution of values ranged from 0.46 to 0.73 in the 90 cases considered. Therefore we correlated team wins with the percentage of home wins and we found a negative correlation of -0.33. This clearly indicated that, in the NBA, worse teams are relatively stronger in home than better teams.

Acknowledging the limitation of not counting with the home/away variable and the strength of schedule, we made two simulations in order to study the possible bias in

[7] We say approximately because some games were played in a neutral venue. From 1950 to 1974, about 1461 games (13.77%) were played in a neutral court (Justin Kubatko from www.basketball-reference.com, personal communication).

our results derived from omitting these variables. Firstly, we analysed the effect of the disparity of calendar on the duality of home/away wins. Recall that all teams in the NBA play the same number of games at home vs. away (currently, 41 games). We used data base of team results of 2006/07, 2007/08 and 2008/09 seasons, downloaded from www.nbastuffer.com[8]. We registered the number of home and away games for each team and we ordered data in function of each team schedule. To achieve simulations, we considered only data from game number 15 to game number 67, which encompasses from about 18 to 82% of season games, because almost of 90% of changes of coaches in the history of NBA occurred within this interval of percentages. We compared the number of home games with the number of expected home games in each partition that would be necessary to maintain the parity of home vs. away games. Note that we considered the number of home games in 54 different partitions (from 15 to 67 games). Then we computed the difference between the observed and the expected number of home games for the 30 teams, for the 3 seasons and for the 54 disparate partitions (30x3x54=4860 cases). Results are showed in Figure 2. The shaded area represents data between the percentiles 5 and 95. Therefore, 90% of data falls under approximately -3 and 3 home games of difference. It is true that there are cases within the 0-5 and 95-100 percentiles where this difference becomes about 6 games (see dashed line), but we may say, with a lot of prudency, that the disparity of calendar do not yield an ostensible difference in home vs. away parity, in particular when data move away from the middle of distribution.

---Figure 2 about here---

[8] This database is a user friendly Excel resource, which allows to program different procedures to analyse the home/field advantage and the strength of schedule. Unfortunately, data are only available from the 2006/07 season.

Secondly, and using the same data base of three NBA regular season games, we analysed differences in strength of schedule in the same 54 partitions. For example, for the 15 games partition, we compared the mean of opponents winning percentage at the end of the season (from game 1 to game 15), with the mean of opponents winning percentage at the end of the season for the 67 remaining games (from game 16 to game 82). In order to avoid scale problems and to count with dispersion of data, we also achieve the same analysis using the coefficient of variation, i.e. the standard deviation divided by the mean. Results are showed in Figure 3. The shaded area represents data between the percentiles 5 and 95. Therefore, 90% of data falls under approximately (-0.05 and 0.05) units of difference. Considering the winning percentage is a [0,1] variable, this represent only between -5% and 5% of variation. It is true that this difference is a little higher for coefficient of variation (dashed line), which indicates that there is also a difference in the dispersion of distributions (some teams play with more homogeneous rivals than others, with regard to the quality of teams). Contrary to the Figure 2, differences become larger in the tails of the graphic.

---Figure 3 about here---

In sum, acknowledging the shortcoming of omitting variables such as home field advantage and strength of schedule, these simulations show that the potential omitted information should no have an important effect on the probability of success. The lack of an available database to facilitate the registration of this great amount of data has been a barrier impossible to overcome. However, after viewing these simulations, we are convinced our results are still robust and they are not importantly biased.

Interpretation and generalization of results

A final commentary regarding our methodology deserves to be highlighted. We have studied all the population of NBA changes within a season in the history of the league. There is no sample of changes, consequently, from a theoretical viewpoint, there is no inference from the sample to the hypothetical population. Therefore, we do not need statistical tests to analyse results[9]. However, this philosophy of analysis excludes any potential extension of results to other populations, such as other leagues, for example. In addition, this excludes any predictive inference about future changes in the NBA. Therefore, we also report statistical tests and confidence intervals in the analysis made with all the population of coaches. Extension to a hypothetical infinite population of coaches' changes would require the assumptions of: (1) the 203 changes of coaches analysed are a random sample of a population of changes in all the leagues of the world; (2) future conditions in the NBA will be the same (there will be no new systematic factors affecting change). The first assumption would be required to respond to the question: Changing a coach within a season improves team performance in basketball (without restriction to the NBA)? And the second assumption would be necessary to make predictions. Taking into account the heterogeneity of basketball leagues in the world, placed in different countries, with disparate sport cultures and systems of competition, the first assumption seems less plausible than the second. Therefore, we should be very cautious for the generalization of results.

[9] We are referring here to the inferential procedure from a sample of changes to the population of changes. We considered the population of changes, because we studied all the mid-season changes occurred in the history of the league. Therefore, we did not need statistical inference for explaining the history of NBA changes of coaches. Note that this is a different case from the procedure to compare the performance of coaches, depicted previously as a statistical comparison between two proportions. In this latter case, it is necessary a statistical test because we compare two samples of games played by two distinct coaches.

Results

A total of 203 changes occurred in the 61 years of history of the NBA (a mean of 3.32 changes by year). Table 1 shows the distribution of changes. As there is heterogeneity in the games played by teams, we normalised the changes per 1230 games, because this is the number of games played in the NBA each season since 2004/05. There is high dispersion in the distribution of changes, but we found a little pattern of association between coach changes and time, because Pearson correlation was -0.12. Therefore there is a small trend of diminishing instability with time.

--- Table 1 about here ---

A total of 158 different coaches were signed. Table 2 shows all the coaches signed and the number of times they were contracted within a season. One hundred and twenty-one of these 158 coaches did not manage any NBA team before, although 6 coaches trained only BAA or ABA teams, and not NBA teams before they were signed. This represents about 76.6% of the coaches and about 59.6% of all the changes made.

--- Table 2 about here ---

Regarding when coaches were replaced, we used the percentage of games played in a season instead of the raw number of games played because of the diversity of seasons considered. Mean was 0.42 and median was 0.41. Therefore the distribution was symmetrical, being 0.23 the first quartile and 0.57 the third. Data ranged from 0.01% (when Buffalo Braves replaced Dolph Shayes by Johnny McCarthy in 1971 after the first game) to 0.96% (when Atlanta Hawks replaced Hubie Brown and signed Mike Fratello in 1981 for playing the three last games of the season). Thirty six of these new coaches did not finish the season either, so they also were replaced within the same season. Twenty-two of them could be considered as "temporary coaches" because they managed four games or less, waiting for the arrival of the coach really desired by team owners.

In Table 3 we show the comparison of performance of the teams managed by the signed coaches compared with the same teams managed by replaced coaches. Results are derived from the use of the binomial comparison, previously explained.

--- Table 3 about here ---

We may say that the vast majority of changes (3 of 4 considering only valid cases) did not significantly influence performance. However, among the 48 significant changes, almost 2 of 3 of these changes improved team performance. Therefore, only about a 15% of coach changes yielded the desired effect. If we take now an inferential approach, we need a statistical test. Using chi-square, we find significant results: χ^2=151.1 (*df.*=2); *p*=0.000, and with the Phi statistic (Grissom and Kim, 2005), we may provide an effect size index: *Phi*=0.89, which can be considered a very high effect. These results reject the hypothesis of equal effect of the three possible outcomes (neutral, worse or better performance). Therefore, it seems clear the prevalence of the neutral effect. We can also modify the analysis creating only two possible outcomes: neutral+worse vs. better, then χ^2=90.75 (*df.*=1); *p*=0.000; *Phi*=0.68, which it is also a very noticeable effect size.

As we show in Table 4, the majority of coaches who got better performance continued the following season in their respective teams (76.67%), but some of them were fired having outperformed the previous coaches (20.00%).

--- Table 4 about here---

We may also simulate the estimate of the percentage of successful coaches in the following years under the assumption of equality of conditions. Table 5 shows these simulations. We used the binomial estimation and the associated 95% confidence interval. Considering that the mean of replacements is 3.32 per year, in the following 10 years it seems that the estimation is very accurate. Obviously, reliability of estimate decreases over time.

--- Table 5 about here---

Once obtained global results about the effect of firing a coach within a season, we were interested to test a model in order to explain success, i.e. which factors increase or decrease the probability of improving performance. We applied the philosophy of analysis of Mayo (1996) and Spanos (2007; 2010), in order to achieve an inductive process of learning from data, using a statistical model and testing its assumptions.The idea behind this approach is that to secure the reliability of any inductive inference one should validate the underlying inductive premises by probing for all possible errors. Model should account with data regularities, obtaining a white noise error term, i.e. a random component without systematic contamination. Therefore, reliability of the model is primary addressed by testing the assumptions using miss-specification tests. This criterion prevails against any index of predictive accuracy of the model (e.g. explained variance, classification error, AIC, ROC curve, etc.). Model should account with data regularities, obtaining a white noise error term, i.e. a random component without systematic contamination. Therefore, reliability of the model is primary addressed by testing the model assumptions using miss-specification tests. Only if assumptions about the error term are met among disparate models, then model selection may be made using the best predictive accuracy.

As we used logistic regression, this is a low-demanding method regarding assumptions (Tabachnick and Fidell, 2007). This technique assumes linearity in the logit and independence of errors. For studying the first assumption, we used graphical representation of the logit of the expected probabilities against each predictor, instead of statistical techniques such as the Box-Tidwell approach (see Tabachnick and Fidell, 2007) because of the presence of zeros and the impossibility of computing the natural logarithm[10]. For studying the second assumption, we used the Wald-Wolfowitz (WW) runs test applied to the sign of residuals (see Spanos, 2010).

First of all, we ran a model considering only the 30 cases of success and the 18 cases of worse performance. The aim was clearly analyse the influence of predictors on the probability of success. We considered as predictors X_1, X_2, X_3 and X_4 In addition, as the number of cases was relatively low, we were able to compute the difference of home games played before and after the new coaches were signed. The procedure for getting

[10] These graphical representations are available from authors upon request. All these exploratory analyses supported the linear hypothesis.

this difference was the computation of the expected number of home games for both partitions[11]. We called this variable as X_6.

Results of the maximum likelihood estimation using STATA software are depicted in Table 6. As WW test yielded non-significant results, we provide three indexes of model performance, such as percentage of corrected classified cases, R-square (the square of the correlation between the dependent variable and the model prediction) and the area under the ROC curve. We also show the parameter estimates and the marginal effects evaluated at the mean of each variable. Note that marginal effects are a measure of effect size, and provide information regarding the importance of each predictor.

--- Table 6 about here ---

Model 1 only misclassified three cases (93.83% of classification accuracy). As we hypothesized, both parameters associated to X_3 and X_4 were negative, which indicated that to the extent that the winning percentage of a team increases, the probability of success decreases, and to the extent that the difference between the percentage of games played by a new coach and the percentage of games played by an old coach increases, then also decreases the probability of success. This latter fact indicates that firing a coach "as soon as possible" increases the probability of success. In addition, this variable yielded the stronger marginal effect. It is highly remarkable that X_6 yielded a negligible effect on the probability of success, which indicates that the

[11] Some games were played in a neutral venue. This fact was taken into account in order to compute the expected number of home games played.

small discrepancies of the home games played by replaced and new coaches do not cause variation in the dependent variable.

Both experience as coaches and former NBA players had a positive effect on the probability of success. Note that we ran several models using the three different indicators of experience as players X_2. The most simplistic indicators (category of player, and number of season played) performed equally well as the sophisticated index created from minutes played and winshares[12] (see Model 2). We decided to select the model containing the number of seasons variable (Model 1), because the player category would be more difficult to code for predictive analysis. Note that players become Hall of Fame members several years after they left the court, so a former player could pilot a NBA team before being considered member of this selected group of historical players.

This first analysis served to support the selection of our independent variables and to identify factors discriminating between unsuccessful and successful decisions. Therefore, we think it is valuable for understanding factors affecting successful performance. However, the reality is somewhat different, because there is a great amount of replacements which have not been considered. Consequently, we were losing information (together with statistical power), with such small sample, and thus we considered in the subsequent analysis all of the 192 valid cases.

Therefore, we tested a new model (Model 3), grouping as "unsuccessful performance" such cases where the performance was worse (18 cases) and neutral (144

[12] Note that four cases were missing, because of the impossibility of obtaining full data about four of these players.

cases). Two reasons justified that decision: (1) although a new coach gets equal performance than a replaced coach, owners of teams lose money, because of a settlement they have to pay to the replaced coach and the money they have to pay to the new coach. Therefore, neutral effects in performance can also be considered as unsuccessful outcomes; (2) the small number of cases (18) of worse performance would not support the implementation of an ordered logistic regression, using three categories for the independent variable (worse, neutral, better).

Results are also showed in Table 6. Note that we did not consider the variable X_6 in the model. The negligible effect found in the first analysis and results of the simulation explained in the previous section support this decision. In addition, we chose the number of seasons played as NBA players as indicator of former player experience. Model correctly classifies 90.10% of the cases. This is not a very noticeable improvement to the naïve classifier (model without predictors), because it would obtain 84.4%, although it is significantly higher[13]. Again, both parameters associated to X_3 and X_4 were negative, but this time marginal effect of X_3 was higher than the effect of X_4. The experience of new coaches both as coaches and former players slightly influenced the probability of success. We may say that results were in line with those obtained in the first analysis, but with some differences; (1) the marginal effects of X_1 and X_2 decreased; (2) the strongest influence on the probability of success is made by X_3, not X_4. Therefore, the experience of coaches is almost irrelevant for increasing the probability of success. In addition, firing a coach in the early stages of the season increases the probability of success, but the effect is also small, because the highest

[13] We ran a model with only an intercept, restricting all coefficients to zero, and yielded significant results, which indicated that predictors are relevant for explaining the dependent variable.

effect is achieved by the winning percentage variable, i.e. to the extent that the winning percentage of a team increases, the probability of a new coach improves performance decreases.

As we explained previously, we also explored the effect of X_5 (experience of coaches in their whole career) in all the model tested, and no significant results were found. The inclusion of this variable did not improve the classification accuracy of the models. Therefore, and considering that this information is not available at the moment of signing a new coach, decision makers are not harmed by do not know it[14].

Finally, and using Model 3 (the whole valid cases), we may deepen into the interpretation of the marginal effects, in order to provide specific examples of the effect of one variable on the probability of success, obviously, ceteris paribus. We evaluated marginal effects of each predictor at different values, and using the mean value as the reference point in the other predictors. A guide about how to achieve this procedure can be found in Wooldridge (2003). Table 7 shows some of these marginal effects. If teams hire a coach with a great level of experience (both as coach and as a former player), then probability of success increases in a non-linear form, but this probability decrease in a more powerful way if the winning percentage of the team increases. Therefore, hiring a coach with high degree of NBA experience increases the probability of success, but this effect is higher if the winning percentage of the team is low and if the change is produced if the new coach has (potentially) much more games to play than the coach he replaced.

[14] We also tested several models using the number of minutes and winshares variable, but did not outperform the classification rate of Model 3. Other specifications, including quadratic and interaction effects were tested, but did not improve the original model.

--- Table 7 about here ---

Discussion, limitations and further research

In this research we made a complete review of the mid-season changes of coaches occurred along the NBA history. We statistically compared the winning percentage of each team at the moment of change, with the winning percentage got by a new coach, in order to obtain a reliable test considering performance achieved in a sample of the whole games of a season. In addition we tested a model in order to explain the successes cases, i.e. when a new coach got statistically better results than the coach he replaced.

Results show that changing a coach only improved performance in about 15% of cases, and decreased performance in about 10% of cases, being the neutral effect prevalent in the majority of cases (about 75%). This result indicates that NBA owners and general managers made bad decisions about 85% of times, because of the financial costs associated to the change. Although we have analysed if new hired coaches continued training in the subsequent season, we have not studied if results obtained in following seasons were the expected by owners. Therefore, our results are restricted to the short-term, i.e. performance within a season[15]. Consequently, we acknowledge that these decisions might be successful in the long-term. Further research should explore this issue, in a similar way as Giambatista (2004) did. Nevertheless, it seems that owners followed a rational thinking with successful coaches, because more than 75% of

[15] Recall that we did not consider play-off games, just regular season games. A new coach could achieve a great performance in the play-off series, but we excluded these kind of games because we did not a similar series of games managed by replaced coaches.

coaches who got better results continued the next season. In addition, as these changes of coaches are a disrupt in team stability, this result could partially support the findings of Montanari, Silvestre and Gallo (2008), who found that team stability and longevity of team relationships have a positive impact on performance.

The probability of success (i.e. the probability of new coaches improves performance), increased with the experience of new coaches (defined by the number of games managed and the number of wins got till the moment of change), and also with the experience as former NBA players (defined by the number of seasons played), but it was inversely related to the winning percentage of teams, and the difference between games played by old and new coaches.

Marginal effects indicated that the experience as a former NBA player had the smallest effect on the probability of success. Although Goodall, Kahn and Oswald (2010) found a strong predictor of coach success, we did not find such effect size. It is important to note that our definition of coach success is different from these authors, because we only restrict success to the improvement of winning percentage within one season.

It seems clear that if owners want to improve the probability of success when they fire a coach, then they have to hire a new coach with high experience as NBA coach and as NBA player. However, these type of coaches use to be the best paid, so owners and general managers have to make a difficult decision, because of the trade-off between experience and salary. Note that hiring a coach with the highest NBA experience, only would increase the probability of success about 15% with respect to a rookie coach, and hiring a coach with a great career as a player only would increase probability about 5% with respect to a coach without experience as a former NBA player. The difference between salaries of both types of coaches could be about $3-5 millions, i.e. a very important difference.

Changing a coach in the earlier period of the season helps to improve performance. Again the effect size is small. This indicates that when things go bad, it would be a wiser decision to fire a coach as soon as possible than to wait for playing more games, because there would not have time enough to re-drive the situation. However, the most important variable in the model is the winning percentage of teams at the moment of change. Worse teams are much more probably to get success than better teams. In fact, the mean of the winning percentage of teams who decreased performance was 0.38 against 0.21 for teams who improved results. Therefore, changing a coach for the better teams of the NBA within a season is a very risky decision, with very low probability of success.

We recognize all these conclusions are highly dependent of the procedure we achieved to compare the performance of coaches. We based our reasoning in the finite population approach for comparing proportions. Note that this procedure requires the assumption of the number of games managed by coaches is a random sample of the hypothetical 82 games that they would play in the whole season. In addition, any sample realization, i.e. any result recorded should be independent from the remaining games. Obviously, this is not exactly the reality. Games played by coaches could be considered a pseudo-random sample of games. In fact, they are not previously determined before a season starts. Although schedule is known, it would be impossible to a priori know how many games would play a fired coach and a hired coach, neither the moment of the season where change would make. Therefore, we believe the random sampling assumption is not severely problematic. Regarding independence, Arkes and Martínez (2010) found the existence of momentum in NBA, i.e. results of games partially depends on results of previous games, once controlled for several factors. It is true that marginal effect of momentum is small (about 3-4%), but this would indicate that the independence assumption would be violated. We think that, considering the small effect size of the momentum variable, our method is robust against this slight departure of independency of observations, but we acknowledge this fact as a limitation of our work.

We also acknowledge that we based the statistical comparison of proportions in the classical normal approximation. However, other methods are available, such as bootstrap for finite populations (see Lombardía, González-Manteiga & Prada-Sánchez, 2004, for a review). Finite sample distribution of the test statistic could be obtained by a parametric bootstrap, and might give different critical values from the normal approximation. Nevertheless, bootstrap techniques have also limitations, and results are dependent on several factors such as the number of replications, the size of the population and the sample, or the different forms of achieving the resampling procedure. One option for further research is to only consider cases when a triangulation of statistical procedures, (such as some bootstrapping methods and the binomial test), agree. However, probably this form of analysis would yield some inconsistent results and several cases should be deleted from the analysis. Anyway, we think that the simplicity of the procedure we used in this research is an advantage, because it allows simple hand computation, making easier its use for practitioners. In addition, there is a strong theory behind this procedure that supports its feasibility.

The finite population approach to test differences in proportions allows counting with an error which decreases to the extent that the sample approaches population. It is true that some authors (see Grissom & Kim, 2005) recommend some threshold values in the combination of n and p for applying the test that are not fully met in our data, but they do not consider the nature of a finite population, because this recommendation is made in a context of infinite population. We also acknowledge the approximate nature of the computed confidence intervals and the conservative characteristic of the binomial difference test using the normal approximation. In addition, other methodological approaches would be also feasible to apply, such as data envelopment analysis (e.g. Fizel & D'Itri, 1996; 1999; González-Gómez, Picazo-Tadeo & García-Rubio, 2011). However, acknowledging the merit of such sophisticated analysis, several caveats arise from the determination of the inputs. For example, González-Gómez, Picazo-Tadeo & García-Rubio, 2011 use team budgets as input as a proxy of quality of teams. However, as Berri and Schmidt (2010) found, in the NBA, teams pay-roll only explain about 6% of variance in wins. Therefore, it seems that is not a good measure of quality or strength of teams. In addition, as we considered the winning percentage at the moment of change, this variable is implicitly a proxy for measuring the quality of teams.

An alternative form of interpreting the finite population approach would have been to compare the aggregated percentage of wins of new coaches (0.41) against the aggregated percentage of wins of coaches they replaced (0.37), i.e. to achieve an unique test of difference of proportions, instead of the 203 tests (192 valid cases) achieved.

Therefore, and considering 15457 games as the total population of games, 7846 the number of games managed by new coaches, and 6596 the number of games managed by old coaches, then 95% confidence interval around the two proportions would be (0.403 ; 0.418) and (0.366 ; 0.383), i.e. in an aggregated form, new coaches significantly improved the performance of their predecessors. This form of viewing the analysis would overcome the problem of the conservative individual tests, but obviously do not allow for individual analysis of the success of change. However, we think it provides useful information regarding the trend of the change effect.

Nevertheless, a major limitation of our study arises for not controlling for mid-season transactions. These types of transactions could improve (or decline) the quality of teams, and could be a source of systematic noise in our analysis. Although data on transactions were available at www.baskteball-reference.com, we think it would be highly complicated to analyse how these changes in rosters could influence results. Nevertheless, the most important changes in rosters are not usually achieved in mid-season time, but when a season ends. Therefore, the possible bias in our results would be of lesser importance that the potential bias of studies considering longitudinal approaches (e.g. Giambatista, 2004), or analysing the seasons before and after the coach were changed (McTeer, White & Persad, 1995). One of the possible solutions would be considering player talent as a proxy for the quality of the roster, in a similar way as Fizel and D'itri (1999) did. However, how to measure talent of players is a controversial issue in basketball (see Berri & Bradbury, 2010; Berri & Schmidt, 2010; Berri, Schmidt & Brook, 2006), and the form Fizel and D'itri (1999) addressed this question is very debatable. Consequently, it is a challenge for further research to try to count with this factor using the best measurement instrument.

In sum, and acknowledging the commented limitations, our research has advanced in the understanding of the effect of hiring new coaches on performance of NBA teams, when change is achieved in the middle of the season. Changing a coach is an important managerial decision which does not guarantee improvement in the short term, i.e. in the same season. Only about 15% of new coaches outperformed in a significantly way their predecessors. In order to maximize the probability of success, highly experienced coaches, with a long career as former NBA players should be signed. In addition, change should be made before season advances. These three factors slightly contribute to increase the probability of success, being the most important variable the winning percentage of the team at the moment of change. Worse teams are more probable to be successful in their change than better teams.

Acknowledgements

We are in debt with www.basketball-reference.com, and especially with Justin Kubatko, who help to correct some inconsistent data. We also thank Serhat Ugur, from www.nbastuffer.com, for his assistance to obtain data. In addition, thanks to Crow, from the APBRmetrics site, for providing some relevant information regarding coaches' salaries. Finally, thanks to John Fizel from Penn University, and particularly to Aris Spanos, from Virginia Tech University, Francisco González-Gómez from Universidad de Granada and Ruud Koning from University of Groningen, for their comments and suggestions.

Appendix: Successful coaches

Coach	Rookie	Team	Season	WP% old coach	Games old coach	WP% new coach	Games new coach	CI-	CI+
Red Auberbach	Yes	Blackhawks	1949-50	0.14	7	0.49	57	-0.62	-0.08
Jack Smiley	Yes	Waterloo Hawks	1949-50	0.23	35	0.41	27	-0.35	-0.01
Clair Bee	Yes	Balltimore Bullets	1952-53	0.00	3	0.24	67	-0.26	-0.22
Dave DeBusschere	Yes	Detroit Pistons	1964-65	0.18	11	0.42	69	-0.46	-0.01
Red Holzman	No	New York Knicks	1967-68	0.41	37	0.62	45	-0.37	-0.06
Phil Johnson	Yes	Kansas City Omaha Kings	1973-74	0.24	25	0.47	57	-0.39	-0.07
Butch Van Breda	No	New Orleans Jazz	1974-75	0.06	16	0.33	66	-0.39	-0.15
Don Nelson	Yes	Milwaukee Bucks	1976-77	0.17	18	0.42	64	-0.42	-0.09
Lenny Wilkens	No	Seattle Supersonics	1977-78	0.23	22	0.70	60	-0.64	-0.31
Dave Cowens	Yes	Boston Celtics	1978-79	0.14	14	0.40	68	-0.43	-0.07
Phil Johnson	No	Kansas City Kings	1984-85	0.11	9	0.41	73	-0.51	-0.09
Wes Unseld	Yes	Washington Bullets	1987-88	0.30	27	0.55	55	-0.41	-0.09
George Irvine	No	Indiana Pacers	1988-89	0.00	9	0.30	20	-0.48	-0.12
Dick Versace	Yes	Indiana Pacers	1988-89	0.21	29	0.42	53	-0.35	-0.06
Bob Hill	No	Indiana Pacers	1990-91	0.36	25	0.56	57	-0.38	-0.03
Larry Brown	No	L. A. Clippers	1991-92	0.47	47	0.66	35	-0.34	-0.04
Gar Heard	Yes	Dallas Mavericks	1992-93	0.07	29	0.17	53	-0.20	0.00
Bernie Bickerstaff	No	Denver Nuggets	1994-95	0.42	50	0.63	32	-0.36	-0.05
Danny Ainge	Yes	Phoenix	1996-97	0.00	8	0.54	74	-0.58	-0.50
Bernie Bickerstaff	No	Washington Bullets	1996-97	0.47	47	0.63	35	-0.32	-0.01
Don Casey	No	New Jersey Nets	1998-99	0.15	20	0.43	30	-0.45	-0.11
Paul Silas	No	Charlotte Hornets	1998-99	0.27	15	0.63	35	-0.58	-0.15
Bill Cartwright	Yes	Chicago Bulls	2001-02	0.15	27	0.31	55	-0.29	-0.03
Hubie Brown	No	Memphis Grizzlies	2002-03	0.00	8	0.38	74	-0.41	-0.34
Johnny Davis	No	Orlando Magic	2003-04	0.09	11	0.28	71	-0.36	-0.02
Mike Fratello	No	Memphis Grizzlies	2004-05	0.31	16	0.61	66	-0.51	-0.08
Avery Johnson	Yes	Dallas Mavericks	2004-05	0.66	64	0.89	18	-0.38	-0.09
George Karl	No	Denver Nuggets	2004-05	0.40	42	0.80	40	-0.53	-0.26
Scott Brooks	Yes	Ocklahoma City	2008-09	0.08	13	0.32	69	-0.39	-0.10
Kiki Vandeweghe	Yes	New Jersey Nets	2009-10	0.00	18	0.19	64	-0.23	-0.14

References

Arkes, J., & Martínez, J. A. (2010). Finally, evidence for a momentum effect in the NBA. *Journal of Quantitative Analysis in Sports*. Manuscript submitted for publication.

Barros, C. P. Frick, B., & Passos, J. (2009). Coaching for survival: the hazards of head coachcareers in the German "Bundesliga"'. *Applied Economics 41* (25), 3303-3311

Berri, D. J., & Bradbury, J. C. (2010). Working in the land of metricians. *Journal of Sports Economics. 11* (1), 29-47.

Berri, D. J., Schmidt, M. B., & Brook, S. L. (2006). *The wages of wins: Taking measure of the many myths in modern sport*. Palo Alto,CA: Stanford University Press.

Berri, D. J., & Schmidt, M. B. (2010). *Stumbling on wins: Two economists expose the pitfalls on the road to victory in professional sports*. FT Press.

Brown, M.C. (1982). Administrative succession and organizational performance: The succession effect. *Administrative Science Quarterly 27,* 1–16.

Bruinshoofd, A., & ter Wel, B. (2003). Manager to go? Performance dips reconsidered with evidence from Dutch football. *European Journal of Operational Research 148*, 233–246.

Fizel, J.L., & D'itri, M. (1996). Estimating managerial efficiency: the case of college basketball coaches. *Journal of Sport Management, 10*, 435–445.

Fizel, J. L. & D'itri, M. (1999). Firing and hiring of managers: does efficiency matter? *Journal of Management, 25,* 567-585.

Frick, B., Pestana, C., & Prinz, J. (2010). Analysing head coach dismissals in the German ''Bundesliga" with a mixed logit approach. *European Journal of Operational Research 200*, 151-159.

Gamson, W., & Scotch, N. (1964). Scapegoating in baseball. *American Journal of Sociology 70*, 69–72.

García, A. (2010, April, 11). Influence of quality coaching shouldn't be overlooked. Downloaded from http://www.nba.com/2010/news/features/art_garcia/04/11/coaching/index.html

Giambatista, R.C. (2004). Jumping through hoops: A longitudinal study leader life cycles in the NBA. *The Leadership Quarterly*, 15, 607-624.

González-Gómez, F., Picazo-Tadeo, A., & García-Rubio, M. A. (2011). The impact of a mid-season change of manager on sporting performance. *Sport, Business and Management: An International Journal*, 1.

Goodall, A. H., Kahn, L. M., & Oswald, A. J. (2010). Why do leaders matter? A study of expert knowledge in a superstar setting. *Journal of Economic Behavior & Organization*, doi:10.1016/j.jebo.2010.11.00.

Grissom, R. J., & Kim, J. J. (2005). *Effect sizes for research: A broach practical approach.* Mahwah, NJ: Erlbaum

Horowitz, I. (1994) On the manager as principal clerk. *Managerial and Decision Economics 15*, 413–419.

Huang, T. K., Weng, R. C., & Lin, C. J. (2006). Generalized Bradley-Terry models and multi-class probability estimates. *Journal of Machine Learning Research 7*, 85–115

Kahn, L. M. (2006). Race, performance, pay and retention among National Basketball Association head coaches. *Journal of Sports Economics 7* (2), 119-149.

Koning, R. H. (2003). An econometric evaluation of the effect of firing a coach on team performance. *Applied Economics 35*, 555-564

Levy, P. S., & Lemeshow, S. (1999). *Sampling of populations: Methods and applications* (3rd ed.). New York: John Wiley & Sons.

Lombardía, M. J.; González-Manteiga, W., & Prada-Sánchez, J. M. (2004). Review of distribution function estimation and bootstrap resampling methods in finite population. [Una revisión de la estimación de la función de distribución y métodos de remuestreo bootstrap en poblaciones finitas]. *Estadística Española, 46* (155), 149-175.

Mayo, D. G. (1996). *Error and the Growth of Experimental Knowledge.* University of Chicago Press.

McTeer, W., White, P.G., & Persad, S. (1995). Manager/coach mid-season replacement and team performance in professional team sport. *Journal of Sport Behaviour, 18*, 58–68.

Montanari, F., Silvestri, G., & Gallo, E. (2008). Team performance between change and stability: The case of Italian Serie A. *Journal of Sport Management, 22*, 701-716.

Salomo, S., Teichmann, K., & Albrechts, C. (2000). The relationship of performance and managerial succession in the German Premier Football League. *European Journal of Sport Management 7*, 99-117.

Schmueli, G., Patel, N. R., & Bruce, P. C. (2007). *Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Microsoft Office Excel with XLMiner*. New Jersey: Wiley

Scully, G. W. (1995). *The Market Structure of Sports*. University of Chicago Press: Chicago.

Spanos, A. (2007). Curve fitting, the reliability of inductive inference, and the error-statistical approach. *Philosophy of Science, 74*, 1046–1066.

Spanos, A. (2010). The discovery of Argon: A case for learning from data? *Philosophy of Science, 77* (3), 359-380.

Steiger, J. H., & Fouladi, R. T. (1992). R2:Acomputer program for interval estimation, power calculation and hypothesis testing for the squared multiple correlation. *Behavior Research Methods, Instruments, and Computers, 4,* 581-582.

Tena, J. D., & Forrest, D. (2007). Within-season dismissal of football coaches: Statistical analysis of causes and consequences. *European Journal of Operational Research*, *181*, 362-373.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Fifth edition. Pearson International Edition.

Van Dalen, H. P. (1994) Loont het om een voetbaltrainer te ontslaan? *Economisch Statistische Berichten 79*, 1089–1092

Van Ripper, T. (2010, May 21). The highest-paid coaches in sports. Downloaded from http://www.forbes.com/2010/05/21/phil-jackson-belichick-mike-shanahan-business-sports-coaches.html

Wooldridge, J. M. (2003). *Introductory Econometrics: A Modern Approach* (Second Edition). South Western

Table 1. Changes of coaches by season

Season	Teams	Avg. games	Coach changes	Total games	Changes x 1230 games	Season	Teams	Avg. games	Coach changes	Total games	Changes x 1230 games
1949-50	17	66	4	561	8.8	1980-81	23	82	4	943	5.2
1950-51	11	64	4	354	13.9	1981-82	23	82	7	943	9.1
1951-52	10	66	1	330	3.7	1982-83	23	82	1	943	1.3
1952-53	10	70	1	351	3.5	1983-84	23	82	1	943	1.3
1953-54	9	72	1	324	3.8	1984-85	23	82	2	943	2.6
1954-55	9	72	0	288	0.0	1985-86	23	82	2	943	2.6
1955-56	8	72	1	288	4.3	1986-87	23	82	3	943	3.9
1956-57	8	72	3	288	12.8	1987-88	23	82	6	943	7.8
1957-58	8	72	2	288	8.5	1988-89	25	82	6	1025	7.2
1958-59	8	72	2	288	8.5	1989-90	27	82	3	1107	3.3
1959-60	8	75	3	300	12.3	1990-91	27	82	2	1107	2.2
1960-61	8	79	0	316	0.0	1991-92	27	82	8	1107	8.9
1961-62	9	80	3	360	10.3	1992-93	27	82	5	1107	5.6
1962-63	9	80	1	360	3.4	1993-94	27	82	2	1107	2.2
1963-64	9	80	0	360	0.0	1994-95	27	82	4	1107	4.4
1964-65	9	80	3	360	10.3	1995-96	29	82	3	1189	3.1
1965-66	9	80	1	360	3.4	1996-97	29	82	8	1189	8.3
1966-67	10	81	3	405	9.1	1997-98	29	82	3	1189	3.1
1967-68	12	82	1	492	2.5	1998-99	29	50	5	725	8.5
1968-69	14	82	1	574	2.1	1999-00	29	82	6	1189	6.2
1969-70	14	82	3	574	6.4	2000-01	29	82	2	1189	2.1
1970-71	17	82	0	697	0.0	2001-02	29	82	6	1189	6.2
1971-72	17	82	4	697	7.1	2002-03	29	82	4	1189	4.1
1972-73	17	82	5	697	8.8	2003-04	29	82	9	1189	9.3
1973-74	17	82	2	697	3.5	2004-05	30	82	10	1230	10.0
1974-75	18	82	2	738	3.3	2005-06	30	82	3	1230	3.0
1975-76	18	82	2	738	3.3	2006-07	30	82	3	1230	3.0
1976-77	22	82	4	902	5.5	2007-08	30	82	2	1230	2.0
1977-78	22	82	5	902	6.8	2008-09	30	82	9	1230	9.0
1978-79	22	82	4	902	5.5	2009-10	30	82	4	1230	4.0
1979-80	22	82	4	902	5.5	Total			203	48521	

Table 2. Coaches signed

1 time	Alvin Gentry, Bob Bass, Bob Hill, Bob Mackinnon, Gene Littles, Lionel Hollins, Phil Johnson, Red Holzman, Richie Adubato
2 times	Alex Hannun, Bernie Bickerstaff, Bill Bertka, Bob Kloppenburg, Bob Pettit, Dick McGuire, Dick Motta, Don Casey, Don Chaney, Don Nelson, Elgin Baylor, Frank Hamblen, George Irvine, George Karl, Herb Williams, Jerry Colangelo, Jerry Reynolds, Johnny Davis, Kevin Loughery, Kevin McHale, Lenny Wilkens, Mike Fratello, Pat Riley, Paul Seymour, Pete Myers, Rex Hughes, Scott Skiles
3 times	Al Attles, Andrew Levane, Avery Johnson, Bill Berry, Bill Blair, Bill Carwright, Bill Musselman, Billy Cunningham, Bob Kauffman, Bob Lanier, Bob Staak, Brendan Malone, Brian Winters, Bucky Buckwalter, Buddy Jeannette, Bumper Hormohlen, Butch Carter, Butch Van Breda Kolff, Carl Braun, Chick Reiser, Chris Ford, Chris Jent, Chuck Dayly, Clair Bee, Cotton Fitzsimmons, Danny Ainge, Darrel Walker, Dave Cowens, Dave DeBusschere, Dennis Johnson, Dick Van Arsdale, Dick Versace, Don Delaney, Donnie Butcher, Donnie Walsh, Doug Moe, Doxie Moore, Draff Young, Earl Lloyd, Ed Gregory, Ed MaCauley, Ed Tapscott, Eddie Jordan, Flip Sanders, Frank Johnson, Frank Layden, Fred Carter, Gar Heard, Garry St. Jean, Gene Sue, Gregg Popovich, Harry Gallatin, Herb Brown, Hubie Brown, Ike Duffey, Jack Smiley, Jay Triano, Jeff Bower, Jeff Van Gundy, Jerry Sloan, Jim Boylan, Jim Brovelli, Jim Lynam, Jim O'Brien, Jim Pollard, Jim Todd, Joe Mullaney, John Carrol, John Kundla, John Logan, John Lucas, John MacLeod, Johnny Egan, Johnny McCarthy, Jonny Bach, Keith Smart, Kenny Natt, Kevin Pritchard, Kiki Vandeweghe, Kim Hughes, Kurt Rambis, Larry Brown, Larry Krystkowiak, Larry Staverman, Lawrence Frank, Mack Calvin, Magic Johnson, Mel Daniels, Michael Cooper, Mike D'Antoni, Mike Evans, Mike Todorovich, Nate McMillan, Paul Silas, Paul Westhead, Randy Wittman, Ray Scott, Red Auberbach, Red Rocha, Richie Guerin, Rick Adelman, Rod Thorn, Rudy Tomjanovich, Scott Brooks, Scotty Robertson, Sidney Lowe, Slater Martin, Slick Leonard, Stu Inman, Stu Jackson, Terry Dischinger, Terry Scotts, Tom Barrise, Tom Marshall, Tom Sanders, Tony Barone, Tony DiLeo, Vince Boryla, Wally Jones, Walt Budko, Wes Unseld, Willis Reed.

Table 3. Performance comparison between new and replaced coaches.

	Number	%	% (valid cases)
Neutral effect	144	70.93	75.00
Worse performance	18	8.86	9.37
Better performance	30	14.77	15.62
No valid*	11	5.41	

*No valid are the cases when the old or the new coach played only one game, so statistical test was not feasible.

Table 4. Continuity of coaches

		Worse performance	Better performance
Continued next season	96 (47.29%)	2 (11.11%)	23 (76.67%)
Did not continue next season	72 (35.46%)	8 (44.44%)	6 (20.00%)
Fired the same season	35 (17.24%)	8 (44.44%)	1 (3.33%)
Total	203 (192 valid cases)	18	30

Table 5. Simulation of the estimate of the percentage of successful coaches in the following years

	Population	Sample	95% CI
Current study	203	203	0.148
5 next years projection	220	203	(0.134 ; 0.162)
10 next years projection	236	203	(0.130 ; 0.166)
50 next years projection	369	203	(0.115 ; 0.181)
Extremely large projection	2000	203	(0.102; 0.194)

Table 6. Results of the models estimation

	Model 1	Marg. effects	Model 2	Marg. effects	Model 3	Marg. effects
Classification accuracy	93.75%		95.45%		90.10%	
R-square (95% CI)	0.760 (0.56 ; 0.84)		0.829 (0.67 ; 0.89)		0.342 (0.22 ; 0.44)	
Area under the ROC curve	0.965		0.968		0.842	
WW test	0.16		-0.19		0.38	
Constant	0.213		-0.598		0.001	
X_1	2.472	0.271	2.760*	0.337	0.747	0.056
X_2	0.121	0.013	1.762	0.215	0.059	0.004
X_3	-3.862	-0.424	-3.23	-0.39	-8.924	-0.673
X_4	-6.767*	-0.744*	-7.692*	-0.94	-1.002	-0.075
X_5						
X_6	0.030	0.003	0.019	0.002		

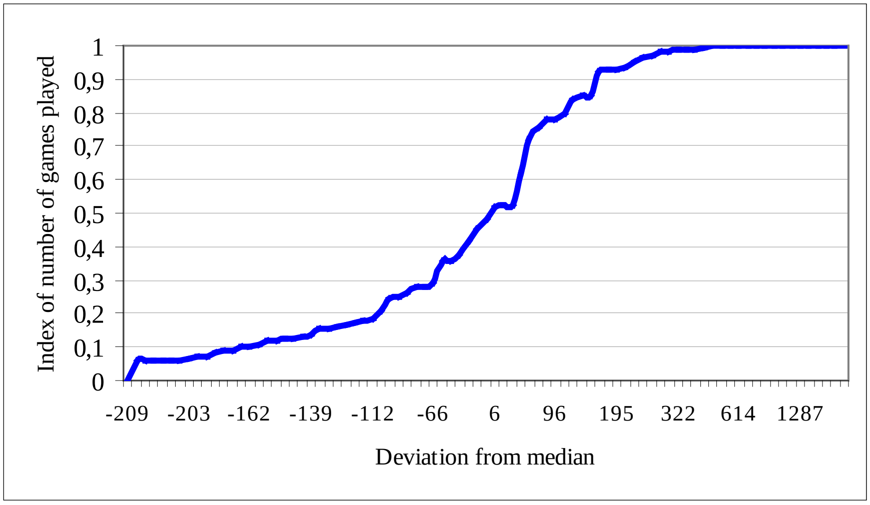
**p<0.05*

Note: To compute confidence intervals around R-square, the R2 program (Steiger & Fouladi, 1992) was used

Table 7. Some marginal effects

Predictor	Change in the value of predictor	Change in the probability of success
X_1	From 0 to 1	5.9%
	From 1 to 2	10.4%
X_2	From 0 to 3	1.9%
	From 5 to 12	3.6%
X_3	From 0.1 to 0.25	-27.7%
	From 0.25 to 0.5	-15.8%
X_4	From –0.25 to 0	-1.9%
	From 0 to 0.5	-7.6%

Figure 1. Index of number of games played after a sigmoid transformation



Note: There are 121 cases in zero.

Figure 2. Simulation of the potential bias derived from omitting the disparity of home games played

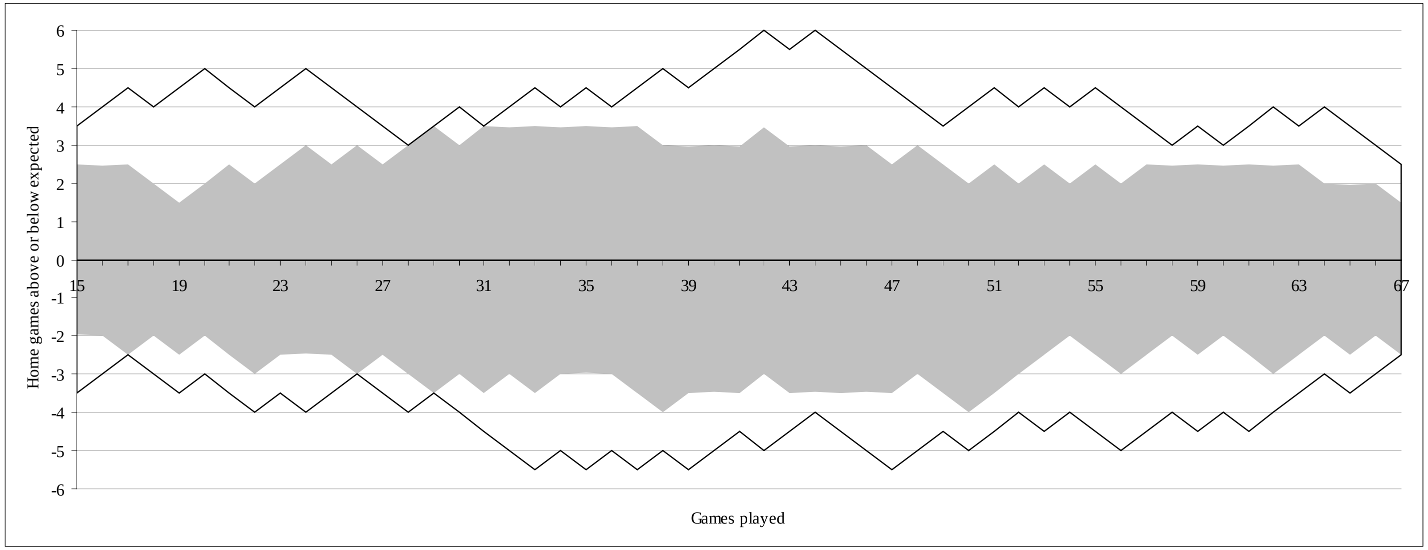


Figure 3. Simulation of the potential bias derived from omitting the strength of schedule