4-16-2013

STATISTICAL EVALUATION AND EFFICIENCY OF MID-SEASON MANAGERIAL TURNOVERS: SIGNS FROM TURKISH SOCCER BETWEEN 1998-2012

Selman Kaldiroglu
Trinity University, skaldiro@trinity.edu

STATISTICAL EVALUATION AND EFFICIENCY OF MID-SEASON MANAGERIAL TURNOVERS: SIGNS FROM TURKISH SOCCER BETWEEN 1998-2012

SELMAN KALDIROGLU

A DEPARTMENT HONORS THESIS SUBMITTED TO THE
DEPARTMENT OF ECONOMICS AT TRINITY UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR GRADUATION WITH DEPARTMENTAL HONORS

DATE APRIL 16, 2013

_____ _____

THESIS ADVISOR DEPARTMENT CHAIR

_____

ASSOCIATE VICE PRESIDENT FOR ACADEMIC AFFAIRS,
CURRICULUM AND STUDENT ISSUES

Honors Thesis

Statistical Evaluation and Efficiency of Mid-Season Soccer Manager Dismissals: Signs from Turkish Soccer between 1998-2012

Selman Kaldiroglu

Advisor: Professor John Huston

March 21, 2013

I. Introduction

For many years, economists have studied the effects of executives on firms' performances. It is a consensus that the CEO of a firm is the main determinant of performance. A manager of a sports club is very much like a CEO. The manager is responsible for achieving the organizational objectives of a club by effective decision-making. Therefore, it is inevitable that a club's performance is directly correlated with its manager in charge. Many academics in the past found that the key cause of managerial dismissals is bad performance (Bruinshoofd and ter Weel, 2003; Balduck and Buelens, 2007; Audas *et al.*, 1999). In this paper, I examine the effectiveness of these critical decisions and determine whether such turnovers contribute to organizational goals of clubs.

The main advantage of using sports as a way to evaluate the effects of manager dismissals is transparency. Unlike many businesses, all soccer clubs have similar organizational structures and goals. There is also less noise in the business and more observable results—points. The points a club earns are better measures for success of the manager than the financial earnings of a firm are (Holmstorm, 1982).

There are multiple theories on terminations of manager contracts. One hypothesis is the "coach effect" also known as the "shock effect" and "honeymoon period" (Wirl and Sagmeister, 2008; Balduck and Buelens, 2007; Hoffler and Slimka, 2003; Brown, 1982). It is the very short-run boost effect once a new manager arrives. It is believed that a new manager will incentivize players to be more enthusiastic in the first few periods. In a way, the players have to prove that they deserve a place in the starting lineup. Therefore they have an incentive to exert more effort, which possibly could result in a better

performance in the next few games. However, as one can imagine, this effect lasts a very short period of time. It is argued that the reverse effect kicks in once the shock effect fades away. The players get used to the new manager as the manager starts to form an idea about the squad therefore ending the honeymoon period. Gamson and Scotch's (1964) approach is divided into three hypotheses; the vicious-cycle hypothesis, the scapegoating hypothesis and the common-sense hypothesis. In short, the vicious cycle hypothesis claims that once a turnover takes place, chaos kicks in and the club performance declines even more in the short-run future. Second, the common-sense theory is the idea that the bad performance is directly related to the manager and, therefore, after a dismissal the club's performance is much better. Finally, the scapegoating theory is maybe the most intriguing one; it defends the argument that firing a manager is a way to explain the bad situation of a club by finding a credible victim to blame. It argues that the board of directors is just looking for a 'reason' for the poor performance that otherwise is ambiguous (Kesner and Sebora, 1994). These theories are discussed in detail later in the paper.

Soccer has been around for many years and each country has developed different soccer cultures. Soccer cultures are differentiated from country to country due to distinctive media climates, importance of soccer for fans and how developed the business side of soccer is in various countries. Turkey, unlike England, Spain, Italy, Germany, has a weak business structure when it comes to soccer. Fans and the media have significant power to manipulate clubs and their decisions. Consequently, relative to leagues such as the English, where soccer is professional and media has not much influence on internal

decisions within clubs, Turkey provides a great environment to test hypotheses to analyze the effectiveness of Boards' decisions to fire managers.

I analyze the presented theories through empirical evidence in the Turkish Super League from 1998 – 2012[1]. Following Koning (2000), Bruinshoofd & ter Weel (2003), and Tena & Forrest (2007), a short and long run analysis are considered separately by modeling in the context of varying time frames. I first talk about past research papers based on this idea and then continue with construction of the data set. Finally, I discuss methods to evaluate the issue. I find evidence for terminations that have no successful recoveries.

II. Literature

Research on this subject is increasing, as soccer becomes more of a business in the modern world. Many academics have done research on how terminations of manager contracts affect performances in specific countries and regions (Salomo and Teichmann, 1999; Koning, 2003; Tena and Forrest, 2007; Bruinshoofd and ter Weel, 2003; De Paola and Scoppa, 2011; Wirl and Sagmeister, 2008; Frick *et al.*, 2010; Balduck and Buelens, 2007). It first started with Grusky (1963) and Allen (1979) with their papers concluding that turnovers result in negative effects in the Major Baseball League in the United States especially if the new coach is from outside the club (and not an existing coach within). Allen (1979) found that there is only an improvement in the subsequent season. Following him, Brown (1982) found that there is no "coach effect" in the NFL. However,

[1] The name was Turkish Super League until 2004. From 2005-2010, it was Turkcell Super League, and from 2011 on the official name became Super Toto Super League.

due to the different hierarchal system of coaching in the NFL, I will not consider his research in the analysis.

The scapegoat hypothesis became an interesting possible explanation for dismissals in the 1980s (Brown, 1982). There have been many papers written and there is much empirical evidence that supports the scapegoat hypothesis (Heuer *et al.*, 2011; Tena & Forrest, 2007, 2012; Audas *et al.*, 2002). Audas and Goddard (2002) and Hughes *et al.* (2010) could not find any evidence of improvement in the English leagues they have analyzed after a manager termination in the short-run. Additionally, in Audas' (2002) study, he found that there is a slight negative impact in managerial turnovers but with an increased variance, which means that the effect could even be positive on some occasions. Similarly, in the Dutch league, the short-run affects are unambiguously negative in the succeeding games after a change of manager (Koning, 2003; Bruinshoofd and ter Weel, 2004). Koning (2003) claims that in the period he investigates, the number of turnovers is notably high that there have to be other reasons that shareholders take such actions. The fans and media are key elements in the decision making process of a manager's future at a club (Audas, 2002; Koning, 2003, Salomo and Teichmann, 2000).

In addition, Salomo and Teichmann (2000) detected experimental evidence that besides media pressure, change in board, especially the president increases the chance of being sacked. These studies by sports economists support the claim that the club's performance, once the manager is sacked, does not improve. In some cases, the performance declines even more or it recovers from poor performance less quickly than teams with similar performances but who have kept their managers, which is what

happened in the Dutch Eredivisie[2] and English Premier League (Dobson and Goddard, 1997; Bruinshoofd and ter Weel 2004). An interesting result came from Cannella and Rowe (2005). They discovered that a positive effect of new coaches only occurs in high rivalry contexts. This could be attributed to the high concentration and hope the new coach brings to the players and their ambition to reach success by 'overruling' the rival team(s). A sub-theory of the scapegoat hypothesis is called 'Regression to the mean' (Salomo and Teichmann 2000; Audas, 2002; Cannella and Rowe, 2005). This result came directly from statistical theory; after extreme outcomes, the succeeding outcomes tend to be closer to the mean. This means that after a dip of performance teams tend to do slightly better in the next few periods. Bruinshoofd and ter Weel (2004) and Audas (2002) controlled for this effect and found negative impact of manager dismissals on team performance. It is also argued that due to the variability of this effect, the individual influence of the manager cannot be evaluated (Gamson and Scotch, 1964). Khanna and Poulsen (1995) argue that the effect is in fact commonplace. They found that firms in financial trouble sack managers more often. In financially troubled clubs, the bad performance cannot be blamed on the manager; other factors such as timely wage payment, or the low quality of the squad due to debt, are significant causes for bad performance. No one would perform well if they did not get paid.

The "Coach Effect" is another theory that has been a center of attention for many economists in the field. The coach effect (shock effect) is a short-run concept, which claims that once a new manager arrives, all the players have a shot at being a regular. Hoffler and Slimka (2003) call this the "honeymoon period" of the team. According to

[2] Dutch First Division

Tena & Forrest's (2007) research paper, there is a small positive effect of manager turnovers in the home games during this period. This is attributed to the eager fans that are looking for a change, supporting their team as much as they can. As one might suspect, not as many fans go to away games and therefore this boost is not visible.

The vicious-cycle hypothesis is another concept that has been researched by various economists. Studies have shown that a manager change mid-season worsens the performance of the team due to information that is lost from the learning process of the previous manager (Greiner, Cummings and Bhambri, 2002). Crossan (1999) claims that the short run improvement is deceiving and that "learning takes time…".

It is crucial to understand different perspectives to construct a control group for this paper. One needs to consider other factors that might affect team performance. Many researchers examined the causes of manager dismissals in depth to see whether there are other factors that influence termination decisions significantly. It has been clear that bad performance increases the chances of a manager's contract being terminated. A British economist, Bachan (2008) performed a study in the English League from 2002 until 2005, supporting results found in early literature. After controlling for coaching ability, it has been observed that being in the relegation zone is much riskier when it comes to terminations, which is a logical and expected outcome (Bachan, 2008; Tena & Forrest, 2007; Audas, 2002). Teams that are in the relegation zone tend to panic and bring change as soon as possible to stay in the division. Several studies that have been conducted on the age, past experience, salary and statistics have shown that high age tends to increase the chance of dismissal while, on the other hand, the experience of coach seems to decrease the likelihood of a dismissal (Hautsch et al, 2001; Salomo and Teichmann,

2000; Porter and Scully, 1982). Another interesting result came from Tena (2007) and Frick *et al.* (2010). In their papers, they found that higher salary and higher budget encourages the shareholders to dismiss managers. This can be credited to the team expectations. A higher paid manager has a higher expected rate of return and a higher budget club is more likely to have higher expectations such as the league title or the UEFA Champions League qualification.[3]

As one can see, the consequences of turnovers mid-season are quite ambiguous due to various studies. Some researchers think it has positive consequences (Fabianic, 1994) while some argue the opposite (Maximiano, 2006). In my study, I do not discuss the causes of turnovers. I instead focus on whether the decision of dismissing a manager is effective in the short run and in the long run, and what affects the ex-post performance of a team once a firing takes place.

III. Data

In this section, I give a brief summary of the relevant data in the subject and how it was collected. I give background information on the league and then continue with basic statistics from the Turkish Super League from 1998 until 2012. It is important to determine how concepts such as *performance*, *success*, and *failure* are defined. Therefore, before any evaluation, all relevant terms are defined. Another key ingredient is the variables considered in the regression. In parts C through F, I focus on what variables were considered and what variables were used in forming a control group to end up with

[3] UEFA Champions League is an annual continental club soccer competition and is known to be the most prestigious tournament in not just European soccer but also the entire world.

a healthier and clearer regression analysis. The short-run model consists of the analysis of a total of eight games (four preceding and four succeeding a termination) whereas the long-run model consists of the entire season. Details of each model are discussed later.

A. League

The *Turkish Super League* is the top soccer division in Turkey. It consists of 18 teams formed from all male players. Each team plays all other teams twice (one away and one home), so there are 34 match days. Each win is awarded with three points, each tie is awarded with one point and a loss is awarded with zero points. At the end of the season, the bottom three teams get relegated to the lower division. The team with the highest number of points wins the title.[4] There have been 30 different teams in the league since 1998. Interestingly, only four teams have won the title within these years. The schedule is similar to a European soccer schedule; starting in August going until January, a two-week break and then ending in early May. Data is available for all the games within this period including all team players, coaches at the time, and points before and after each game and locations.

B. Dismissals

In the past 14 years there have been 4284 games played and 376 total managerial contracts for 30 teams in the Turkish Super League. 190 of these were end of contracts, while 186 of these contracts were either terminated by the club or by both parties (meaning voluntarily or involuntarily)[5]. We do not distinguish between these since voluntary resignations are mostly cases where fans and media provoke resignation. Out

[4] Exception: 2011-2012 was played with the Play-Off system. At the end of the season the top 4 teams are put in a group starting with half of their points in the regular season. Each team plays 6 games (home-away) in a 'mini-league' format.
[5] *Turkish Football Association Official Website:* www.tff.org

of these 186 terminations, there were 117 mid-season, 41 summer and 28 winter turnovers. There have been near 16 mid-season turnovers on average per season for 18 teams every year. Realize how large this number is. The English Premier League average number of mid-season manager dismissals in the last five years is around seven per season for 20 teams[6] (Koning, 2003). The excess nine turnovers bring doubt on the effectiveness of manager sackings. We exclude summer and winter turnovers since the team squads also change during those periods; the direct effect of a new coach is vague due to altered team quality and opponent qualities. We also do not consider end of contracts since they are almost always at the end of seasons.

C. The 'Regression' – Short Run

In this section, I briefly introduce the regression model. In the regression model, the dependent variable Y is the performance after the dismissal because that's the key statistic that determines the value of managerial changes. Alternatively, the difference between the performances post and ante is a likely candidate for the dependent variable, which is explored later in the paper (see Robustness). Only variables that could possibly and easily be collected and interpreted are used in the analysis. These variables are performance before the turnover, club budget, coaching experience, position of team, opponent quality and team average quality. Variables such as old coach ability, new coach ability, location, team expectations are trickier to quantify. Ability of coaches is measured by looking at past experiences and recording successes and failures. Team expectations are measured by looking at the recent history of the club[7]. Precisely, the

[6] *Turkish Football Association Official Website:* www.tff.org

[7] Note that the quality of a team, i.e. the ending ranking in the current year might have been useful however would have caused an endogeneity problem as the dependent

average number of points per game is used as a proxy for ability of coaches. Variables such as changes in board and promoted teams are the dummy variables. They will get a value of one if there are any changes in the board or if the team was promoted into the division that year and will get zero otherwise.

Due to the intricacy of the problem, the starting equation (1) will change over time as each alteration is explained and interpreted in a logical and efficient manner.

$$Y_{xp} = \alpha + \beta_1 X_{xa} + \beta_2 X_{ahome} + \beta_3 X_{bdg} + \beta_4 X_{exprnc} + \beta_5 X_{multpl} + \beta_6 X_{ncabil} + \beta_7 X_{ocabil} + \beta_8 X_{oqante} + \beta_9 X_{oqpost} + \beta_{10} X_{phome} + \beta_{11} X_{pnca} + \beta_{12} X_{tavq} + \beta_{13} X_{tpos} + \beta_{14} X_w + \beta_{15} X_{ptq} + \Upsilon_1 X_{prom} + \Upsilon_2 X_{brd} + \varepsilon \quad (1)$$

D. The 'Regression' – Long Run

I will be running a long-run regression to see whether the expected performance increase occurs after the new manager has a few weeks to adjust. The answer to this question might change the entire outcome of the thesis. If there is an obvious performance increase in the long run but not the short run, this might mean that the termination was effective. However if there is an opposite result, this might support the *coach effect* that was discussed earlier in the introduction. The long run is a more complicated question due to variables that cannot be controlled for (such as 'team expectations'). To fix some problems that arise when looking in the long run, the data was adjusted to include extra variables such as 'games played before' and 'games played after' (see Issues). The initial regression model follows the same intuition but excludes variables that relate only to short run.

---

variable *ex-post points* determines the ending ranking of a team. This was avoided by including the *team expectations* variable instead of *team quality (ending ranking)* without losing much information about the quality of the team.

$$Y_{xpp} = \alpha + \beta_1 X_{xap} + \beta_2 X_{bdg} + \beta_3 X_{brd} + \beta_4 X_{pnca} + \beta_5 X_{multpl} + \beta_6 X_{ncabil} + \beta_7 X_{ocabil} + \beta_8 X_{tavq} + \beta_9 X_w + \beta_{10} X_{tpos} + \beta_{11} X_{exprnc} + \beta_{12} X_{ptq} + \Upsilon_1 X_{prom} + + \Upsilon_2 X_{brd} + \varepsilon \quad (2)$$

E. Measurement of Performance

In order to measure the before and after performances of teams once a termination takes place, I use points as reference points. A win (three points) is an obvious evidence for good performance. Not many managers get fired right after they win a game. A loss, on the other hand, could be a sign of poor performance depending on the quality of the opponent and the team itself. A team trying to survive in the league losing against the titleholder is not a sign of poor performance necessarily. In order to distinguish between poor and good performance, we require a benchmark performance level. Bruinsfhood and ter Weel (2003) offer various options for a benchmark level. Since this benchmark is different for every team—that is, performing poorly is defined differently for every team—we use an objective proxy. Following Bruinsfhood and ter Weel (2003), we consider the past four games of each dismissal in the league to see if there is solid evidence of poor performance. This number should balance the short-run concept and give us a variety of opponents. Then we consider the first four games after the dismissal to see the effects of the new head coach on the team's performance. To do this, the total number of points are collected, recorded within those four games and used for the regression.

Key variables that influence the team's performance are club budget, coaching experience, ability of the old and new coaches, team quality, opponent quality, changes in board, game locations, position of team, team expectations, and salary of old coach. In

order to determine if hiring a new manager is effective, a control group is created. It consists of teams that perform as poorly but take no action whatsoever regarding the manager (see Construction of Control Group). The main disparities are the average points before and after the dismissal date. The control group plays a crucial role determining the effectiveness of managers by illustrating the regression to the mean effect, which is discussed later in the paper.

F. Variables

Some variables will be included in the long run or short run only, whereas some variables will be included in both. The reasons for each variable are explained below. This section discusses why some variables are included in the model and how problems with the variables are fixed throughout the research. In order to best capture the effectiveness of managerial dismissals, it is important to consider each variable and its intuitive importance in the regression. The statistical importance of variables are analyzed and, if certain variables that turn out to be statistically relevant are further examined.

Ex-Post Performance (xp/xap): this variable is the main dependent variable in the model. **xp** is the number of points won in the four games after the dismissal whereas **xap** is the percentage of points won in the rest of the season. **xp** and **xap** are used in the short-run and long-run regression models, respectively.

Ex-Ante Performance (xa/xaa): is included in the short-run and long-run regression equations as independent variable. **xa** and **xaa** are measured similarly as in ex-post performance. Furthermore, the data is used to find the control group-teams that have similar performances but did not dismiss their managers.

Club Budget (bdg): when thinking about successful teams in the past, it seems like teams with high club budgets are successful. Even though there might be a circular reasoning in this argument, the model will determine if there is an effect at all. Thus, we introduce this variable which measures the annual club budget in millions of U.S. Dollars. A higher club budget does not just lead to higher and guaranteed wages for players, which inevitably motivate players more to perform well but also allows teams to attract better players. This is included in the short-run and long-run models.

Experience (exp): records the number of years that the manager worked as a head coach. The experience of the new manager is relevant because highly experienced managers tend to make fewer novice mistakes and have learned certain things over time that can put them into advantageous situations. A great manager with zero experience might make a stupid decision under pressure that might cause poor performance, whereas an experienced manager, from experience, might make a better decision. This variable appears in both models. Furthermore, it is also used to form another independent variable **multiple** (see *multiple* below).

Team Position (tpos): measures the position of the team at the time of the dismissal with one being the top rank. The current team position reflects the most about the teams' ex-ante performance. There is a high correlation between this variable and performance before the dismissal within that season. However, this correlation is not significantly decreasing the significance of the other variables; therefore it is included in the regression to give an idea of the significance of an extra point, thus reflecting how much pressure the team is underneath (see *Issues*). **Team position** makes more sense when used with **Team Expectations**, because team is underperforming only if it is

positioned below its expected position. It is included in both versions and both short-run and long-run regressions.

Ex-Ante Opponent Quality (oqante): is the sum of the average team positions of the opponents, which is a decent index to measure the difficulty of the four-game period. This variable reflects the difficulty of the teams played in the short run model. The lower it is, the harder the fixture of the team. This is included in the model in order to take into account the effect of harder or easier teams on performance. A team with a low expected position might play tough teams, which might not necessarily mean that the team is underperforming. Because this variable only includes teams in the four games, it is not included in the long run model.

Team Expectations (tavq): is the average position of the teams in the last five years of participation in the division[8]. The team expectations variable is a significant variable to include in the model. A team is underperforming only if it is performing below the expected performance.

Old Coach Ability (ocabil): we measure the ability of a coach by averaging points earned per game over his/her entire career. It reflects how much the old coach that was fired can affect the team's performance. A high averaged coach is more likely to have more of a positive contribution on the performance of a team. This variable is included in both short-run and long-run regressions. It appears either alone or as the difference between the old and new coach abilities. Note that, an old coach can have effects on teams that could last even after his dismissal. Such effects include formation,

[8] Promoted teams that might not have participated in the division in all of the five years are measured by taking the average of the years that they have participated.

set pieces, and tactics. And therefore, this effect can be measured by including old coach ability as an independent variable in the model.

Ex-Ante Location (ahome): counts the number of home games in the preceding four games of a team. Location matters highly in the soccer world. Playing in a field that players are used to, in front of their own fans, gives them a confidence that cannot be achieved in games in other stadiums. Usually, teams alternate and play one week home and one week away; however, there are special circumstances that might have two away games in a row or vice versa. This variable is only used in the short-run analysis.

Promoted (prom): is a dummy variable that is assigned a one if the team is promoted from the lower division into the First division the previous year and zero otherwise. Newly promoted teams tend to have less pressure and tend to be more volatile. A highly skilled coach might be able to influence a promoted team less due to low expectations and confidence of the club in a rigorous division. On the other hand, players might over perform in certain situations in order to get the attention of many more scouts that observe the First division and are interested in transferring players.

New coach ability, ex-post Opponent Quality, ex-post home games, week, board, multiple, promoted, promoted team quality, and **promoted new coach ability** are other variables that play crucial roles in the regression analysis. See appendix for definitions and functions of these variables.

F. Other Data

The following data was collected in order to aid the regression in ways that will be discussed later in the paper.

Ex-Post Games (gp): is the number of games played under the new coach. It is only used in the long-run model because for the short-run model this is fixed to be four for all observations. On some occasions, the dismissal takes place late in the season and early on others. This creates a difference in the number of games coached by the new coaches. The high variance associated with a low number of games creates an ambiguity in analyzing the results.

Ex-Ante Games (gp4): analogously, this variable is the number of games played before the dismissal under the old coach.

IV. Model

Now, the control group is formed and discussed. Following, the model is introduced in a detailed manner.

A. Construction of a Control Group

In order to first answer the question of whether firing aids teams, a control group has to be formed. The reason behind this lies in comparing performance changes by employing a method called difference-in-difference[9].

The motive for constructing a control group is to measure the usefulness of a dismissal that occurred mid-season. If teams with who did fire their managers recover stronger than teams who did not fire their managers, then this can be used as direct evidence for the efficacy of mid-season managerial turnovers. Additionally, the control group can be used to determine how much of the improvement in performance is due to

[9] This method is also employed by Bruinshoofd and ter Weel (2003) in their analysis of the Dutch League managerial dismissals.

coaches and how much is due to regression to the mean[10].

Firstly, teams that exhibited similar ex-ante performances as the regression group are selected after a benchmark average point was determined. Observations with similar performance and no sacking are the most straightforward alternative (Bruinshoofd and ter Weel, 2003). The cutoff line between performing poorly and not poorly is decided to be one point per game in a four-game period. The benchmark point is selected by looking at the average points before a dismissal of the regression group observations. Thus, any time a team earns less than four points in any given four consecutive games that period is included as an observation for the control group.[11] Points during the four-game period before and after are used to measure performance after and before. The same observations are also used to get long-run recovery observations. That is, data is collected similar to the long-run data collection for the regression group discussed earlier.

Figure 1 compares the ex-ante performances of the control group and the regression group. As one can see, the control group and the regression group have similar performance levels except for the small number of sackings that occurred even after decent performances[12]. The average points for the control group before a performance dip is 2.98 whereas this number for the regression group is 3.05. Thus one can conclude that similar observations were chosen to compare these two groups.

[10] As defined earlier, regression to the mean is the issue that if a variable takes an extreme value, the surrounding measurements tend to be closer to the average.
[11] Note that any observation that had a managerial turnover was excluded from the control group.
[12] Ex-ante performances with more than four points in the four game period.

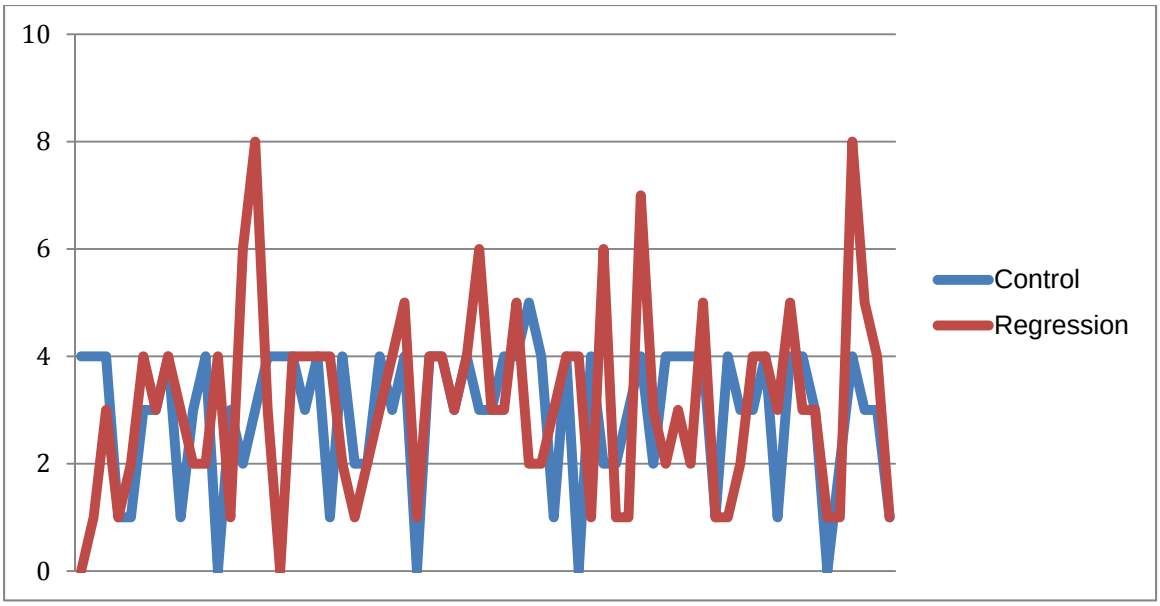


Figure 1. Ex-Ante Performance Data Comparison

B. Effectiveness of Manager Dismissals

The paper continues by explaining how the control group is used to measure the effectiveness of managerial turnovers in the Turkish First Division.

The differences between ex-post and ex-ante performances are recorded for both groups. Then, the difference between the differences, which were recorded, is calculated, hence arriving at our difference-in-difference estimator.

Table 1. *Difference-In-Difference Estimator Descriptive Statistics*

Test Statistic	$(X_p - X_a)$	$(X_p^c - X_a^c)$	$(X_p - X_a) - (X_p^c - X_a^c)$
Coefficient	1.487	1.603	-0.1157
Std Dev.	(0.2767)	(0.3379)	(1.084)
t-statistic	5.374**	4.743**	-0.1067

*** Significant at 1% level*

The first column in *Table 1* depicts the improvement after a dismissal during mid-

season. Indeed, teams perform about 1.5 points better once a dismissal occurs. However, a second glance at the table casts doubt on the cause of the improvement. The second column reveals that control group teams recover even better (about 0.11 more points on average) than teams that actually fired their coaches. The difference-in-difference estimator's t-statistic is calculated to be -0.1067 suggesting no significant difference between groups. This implies that firing managers are not effective and teams who take action should not fire managers hoping for a greater recovery compared to teams who do not take action. This method also takes care of the regression to the mean issue. One might think that part of the improvement is due to the regression to the mean; however, the entire effect in the control group is the regression to the mean effect. A turnover hence would only improve performance if its coefficient were higher than that of the regression to the mean effect, which is not the case.

It should be noted that some firings did lead to superior performance. In the following pages, we seek to discover the characteristics associated with improved performance with a coach replacement.

C. Issues

In this section, we discuss the issues related to the regression models introduced earlier. The initial model constructed by "soccer" intuition is analyzed with econometric and statistical methods in order to have a better understanding of the effects of a new coach. Problems that are notorious in econometrics such as endogeneity and heteroskedasticity as well as various tests are given further attention and resolved in an appropriate fashion. In addition, some variables that *hurt* the regression due to multicollinearity will be considered in the analysis of the model.

Outliers – RStudent & DFBETAS:

Influence statistics are used to discover single observations that are extreme and highly influential to the model and therefore, these observations may be considered to omit in order to not influence the model with only one variable. We use *DFBETAS* and *RStudent* tests that measure the difference that a single observation makes to the regression results. *RStudent* measures the ratio of the residual of the model to an estimated standard deviation, while *DFBETAS* measures the scaled difference in the estimated coefficients between the original equation and the regression line that is estimated without the outlier.

Figure 2. DFBETAS – Board Variable Outliers

One can realize the peculiarity of one single observation in the Figure 2,

above. *DFBETAS* statistical method is made use of in order to discover observation number 48 that had a critical impact on the variable *board.* A further look into the variable indeed confirms the test. The observation includes a dramatic boardroom change, specifically a resignation of the main director of the club right before the dismissal. Since there might be additional factors that cannot be measured in a chaotic period as such, and since it is the only variable with such a change the variable board is omitted from the entire model. [13]

Figure 3. RStudent – Influence Statistics

RStudent test is run to reveal the outliers in the long run model and as one can see in *Figure 3* there exists an outlier in our sample. The RStudent test suggests that observation 87 is significantly different in that its residual is nearly two standard

[13] Notice that omitting the variable and omitting the observation are equivalent since there is only one observation that takes a value of 1 in the sample.

deviations away from zero. This observation was carefully reviewed and there is indeed a peculiarity associated with this observation. The observation included a top team that had European fixtures, cup fixtures as well as league fixtures within the four game period in addition to an investigation from UEFA [14] which might have cumulatively caused stress resulting in strange performance given the independent variables. Therefore, in order to not alter the model based on an outlier, this observation was taken out of the regression.

Heteroskedasticity

In the long-run version of the model, there exists an issue to be resolved before beginning the analysis. When comparing different observations in the long run, there is not a fixed number of games. Certain observations entail fewer games than others depending on the date of the turnover. Some new managers coach for more than 20 games whereas some only for five. Because the dependent variables in this regression are percentages, it is easier to have a high percentage by winning only a few games if the number of *games played* is a small.

Table 2. EViews Breusch-Pagan-Godfrey Test

Heteroskedasticity Test: Breusch-Pagan-Godfrey			
F- Statistic	10.31406	Prob. F(79,36)	0.0017**
Obs*R-Squared	9.624265	Prob. Chi-Square(79)	0.0019**
Scaled Expl. SS	12.82434	Prob. Chi-Square(79)	0.0003**

***Reject Null Hypothesis at 1% level*

To test for heteroskedasticity, the Breusch-Pagan-Godfrey Test was run on the long run model. The null hypothesis, which is the existence of homoskedasticity, is rejected at the five percent significance level (see second column above). This means that

[14] Union of European Football Association

indeed the hypothesis that heteroskedascity exists cannot be rejected.

If the model is sorted in ascending order of *games played*, then the Figure 4 explains the problem. Therefore, the model is weighted inversely with the variable gp^2. One can see the correlation between the residuals in the figure.

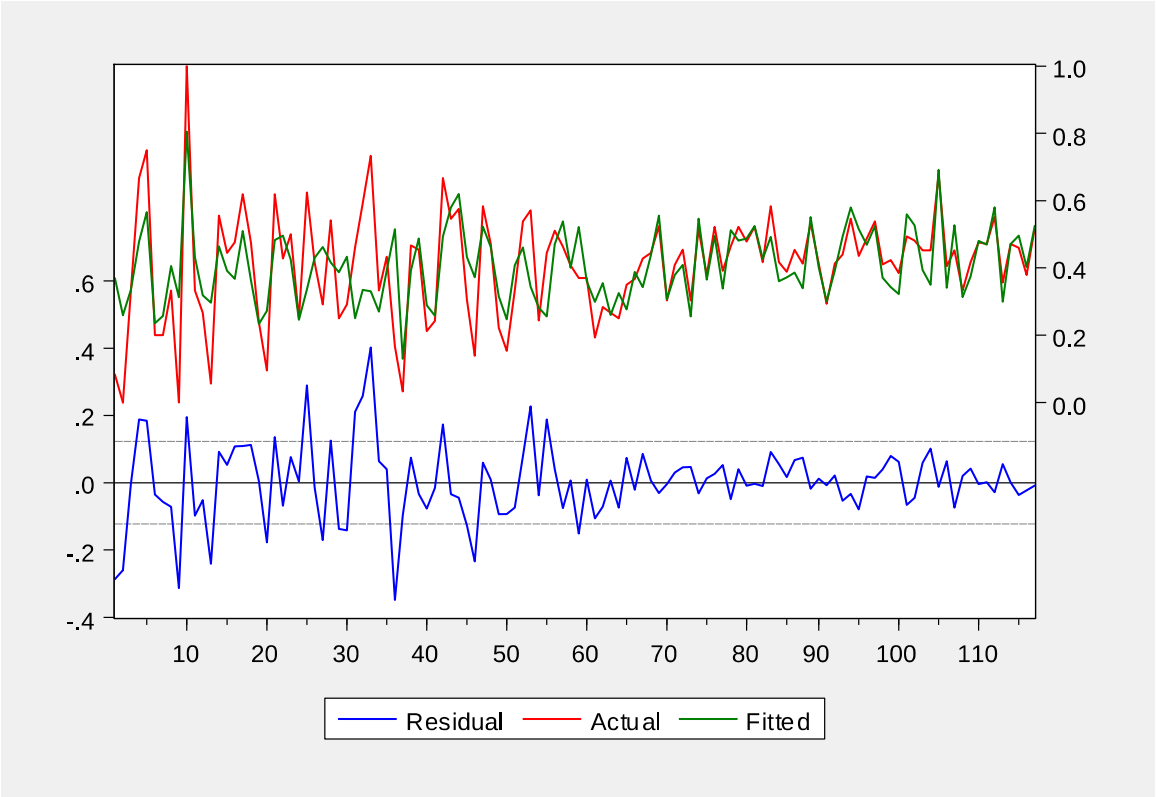


Figure 4. Heteroskedasticity – Residuals of Ordered Observations

As the games played increases, the variance associated with the residuals decreases, strengthening the stability of the model. To understand clearly, one can think of it as two sub-populations (ones with less than 10 games played and ones with more than 10). These two sub-populations have different variances than one another. Once the model is adjusted, the following (*Table 3*) portrays the new model's regression table for the long run.

Table 3. Weighted Long Run Initial Model

Dependent Variable: XPP
Method: Least Squares
Date: 03/06/13 Time: 14:43
Sample: 1 86 88 117
Included observations: 116
Weighting series: GP^2
Weight type: Inverse variance (average scaling)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.393071	0.146568	2.681828	0.0085
XAP	-0.010934	0.115694	-0.094504	0.9249
BDG	2.41E-09	1.85E-09	1.304575	0.1949
PNCA	-0.199882	0.130826	-1.527846	0.1296
MULTPL	0.007110	0.005842	1.217176	0.2263
NCABIL	0.088752	0.059554	1.490283	0.1392
OCABIL	-0.011404	0.034227	-0.333175	0.7397
TAVQ	-0.002436	0.003594	-0.677818	0.4994
W	-0.001048	0.001909	-0.548958	0.5842
TPOS	-0.002498	0.004472	-0.558514	0.5777
EXPRNC	-0.009110	0.008043	-1.132617	0.2600
PROM	0.206040	0.216460	0.951861	0.3434
PTQ	0.000671	0.008744	0.076725	0.9390

Weighted Statistics

R-squared	0.295003	Mean dependent var	0.379419
Adjusted R-squared	0.212868	S.D. dependent var	0.234155
S.E. of regression	0.089115	Akaike info criterion	-1.892507
Sum squared resid	0.817968	Schwarz criterion	-1.583915
Log likelihood	122.7654	Hannan-Quinn criter.	-1.767236
F-statistic	3.591667	Durbin-Watson stat	1.794945
Prob(F-statistic)	0.000173	Weighted mean dep.	0.433950

Unweighted Statistics

R-squared	0.220485	Mean dependent var	0.410333
Adjusted R-squared	0.129668	S.D. dependent var	0.159600
S.E. of regression	0.148894	Sum squared resid	2.283439
Durbin-Watson stat	1.771922		

Elimination

The initial short-run model presented in the earlier sections yield Table 4. The model seems to include variables with low significance and this might be caused by correlated variables or redundant variables that might have been double counted (See *Table 4*).

Table 4. Short Run Initial Model

Dependent Variable: XP
Method: Least Squares
Date: 03/13/13 Time: 12:31
Sample: 1 117
Included observations: 117

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-5.282171	3.650269	-1.447063	0.1510
XA	-0.059613	0.142637	-0.417939	0.6769
PHOME	1.011446	0.473067	2.138061	0.0350**
AHOME	1.081434	0.592422	1.825445	0.0709*
EXPRNC	-0.029847	0.205751	-0.145063	0.8850
MULTPL	0.013145	0.154224	0.085234	0.9322
NCABIL	3.630229	1.165829	3.113860	0.0024***
OCABIL	-1.125871	0.975695	-1.153917	0.2513
OQPOST	0.044433	0.026447	1.680092	0.0961*
OQANTE	0.029270	0.025666	1.140419	0.2569
TPOS	-0.110555	0.084420	-1.309583	0.1934
TAVQ	0.125056	0.091282	1.369998	0.1738
PTQ	-0.402651	0.238485	-1.688366	0.0945*
PROM	5.226471	4.749238	1.100486	0.2738
PNCA	-0.410931	2.611391	-0.157361	0.8753
W	-0.000216	0.028021	-0.007698	0.9939
BDG	2.55E-08	4.90E-08	0.520524	0.6039
BRD	2.083276	2.512256	0.829245	0.4090

R-squared	0.328274	Mean dependent var	4.606838
Adjusted R-squared	0.212927	S.D. dependent var	2.632546
S.E. of regression	2.335519	Akaike info criterion	4.674983
Sum squared resid	540.0101	Schwarz criterion	5.099933
Log likelihood	-255.4865	Hannan-Quinn criter.	4.847508
F-statistic	2.845975	Durbin-Watson stat	1.868411
Prob(F-statistic)	0.000620		

**Significant at 10% level, **Significant at 5% level, ***Significant at 1% level.*

For instance, the p-values of *multpl* and *exprnc* are 0.91 and 0.99 respectively which suggests the insignificance of these variables. However, one should note that the correlation coefficient between these variables is 0.98. Therefore, before concluding anything about these variables we need to check the model by excluding only one of them to see the significance of each individual variable. When deciding what action to take when comparing two correlated variables, one needs to consider the individual

contribution to the regression as well as the insignificance reductions of other variables[15] caused by including the variable in the regression. First, note that *multpl* is defined as the product of *exprnc* and *ncabil.* The variable *ncabil'*s range is around 1 and its variance is low, which explains why the correlation coefficient is quite high. Exclusion of *multpl*, it turns out, does increase the significance of the other variables. Thus, *multpl* was excluded from the model but *exprnc* is still included in the model. One reason of the low significance of these variables might be that even though experience gives an idea of the probability of novice mistakes from the new coach, this aspect might already be reflected in the *ncabil* variable. In addition, *exprnc* and *multpl* were regressed with variables that had near-zero correlation coefficients and similar insignificance was observed. The long-run model was also examined for the significances and correlations for the same variables.

The similar procedure yields the same result; that is, *exprnc* is best left in and *multpl* is best left outside the model and exclusion of *multpl* does more benefit than *harm* to the regression.

[15] We will refer to this effect as "harming" other variables.

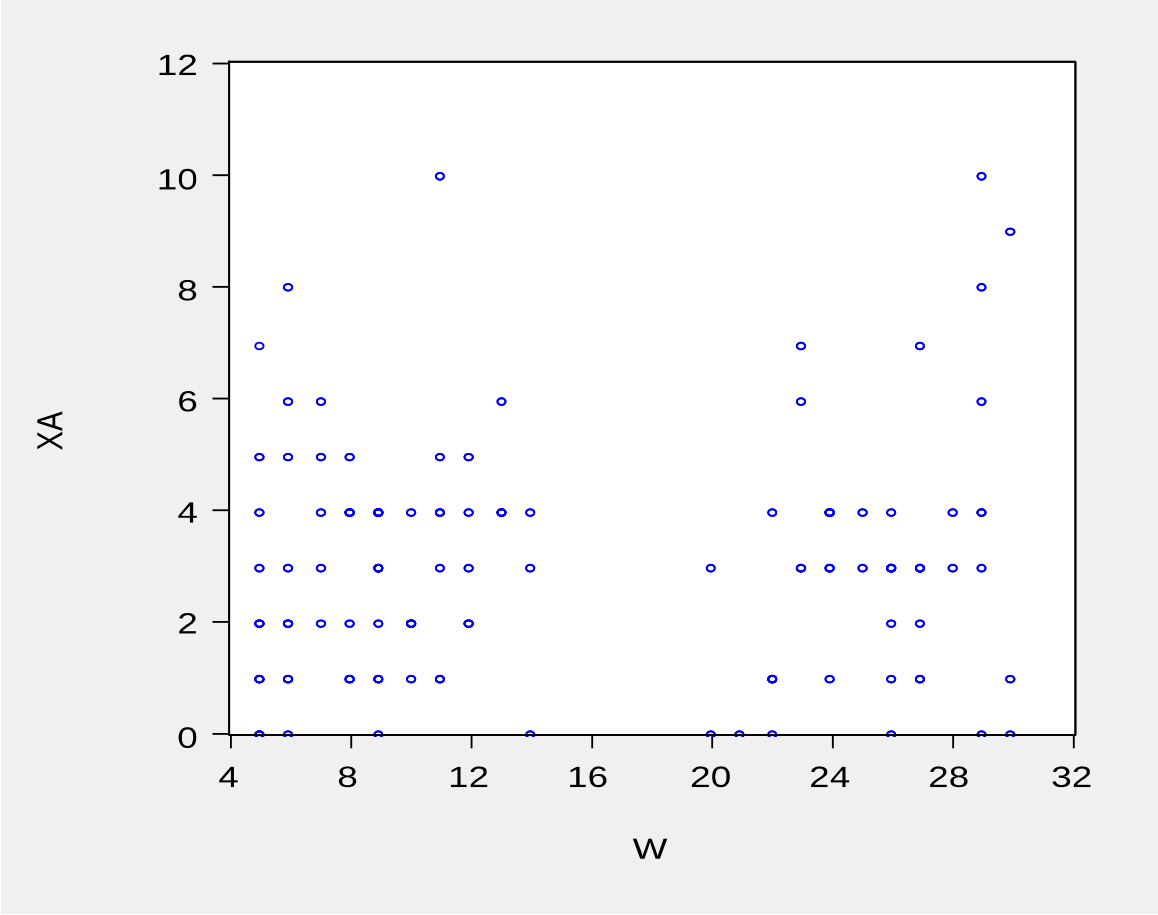


Figure 5. Trend - Ante-Performance vs. Week

The intuitive meaning of the left hand side of our regression equation is the performance under the new coach. Therefore, expecting variables related to ex-ante performance to explain the ex-post performance is not the most logical argument unless there are other reasons we might want to factor in. These variables were included and recorded firstly in order to explain the change of performance and will be used further on in the paper in order to check for robustness. However for now, "ante variables" (*oqante, ahome, xa, xap)* will be excluded in order to avoid any confusion associated with the true meaning and analysis of the regression model. The variable *ocabil* is analyzed within the model and found to be insignificant, even though there was enough reason to believe that it might influence ex-post performance as stated earlier. Similar results can be argued for

xap, *xa* and *oqante.*

The variable that records the matchday of the first game of the new coach, *week (w),* is observed to be consistently insignificant in the long-run model as well as the short-run model. Exclusion of this variable leads to upturns in significances of other variables and inclusion of it contributes indiscernibly to the model. The reason why the week variable was recorded is that turnovers that occur in later weeks are exposed to more pressure due to the time limit whereas turnovers that occur in earlier matchdays of the season have no such worry. The pressure can be realized mostly in the ex-ante performance variable. High pressure increases the likelihood of a firing; thus, it is expected that turnovers that occur in later weeks could have better ex-ante performances than dismissals that occur early on in the season (Audas, 2002; Konig, 2003). The following figure is evidence of the opposite and can be used to logically eliminate this variable from the model. *Figure 5* depicts the relationship between ex-ante performance and the time of the dismissal.

Table 5. Performances based on Matchdays

Ex-Ante Performance (*Xa*)		
Turnover Week[16]	<17 (1st Half of Season)	>17 (2nd Half of Season)
Mean	3.116	3.125
Std. Dev.	1.959413	2.339531

As one can see, there is no trend between the turnovers and the direct performance before a turnover occurs. *Table 5* strengthens the argument by comparing the sample means of performances before and after a given dismissal. The miniscule difference between the means and considerably close standard deviations support our

[16] No observation on the 17th week exists.

claim.

Table 6. Short-Run Edited Model

Dependent Variable: XP
Method: Least Squares
Date: 03/13/13 Time: 20:42
Sample: 1 117
Included observations: 117

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.181193	2.154434	-1.012421	0.3136
PHOME	1.094058	0.458337	2.387018	0.0187**
NCABIL	3.695695	1.007863	3.666862	0.0004**
TPOS	-0.095120	0.069165	-1.375255	0.1719
TAVQ	0.097827	0.081727	1.196996	0.2339
PROM	5.255042	4.470785	1.175418	0.2424
PTQ	-0.418205	0.228571	-1.829650	0.0700*
PNCA	-0.322220	2.401300	-0.134186	0.8935

R-squared	0.267005	Mean dependent var	4.606838
Adjusted R-squared	0.219932	S.D. dependent var	2.632546
S.E. of regression	2.325103	Akaike info criterion	4.591332
Sum squared resid	589.2654	Schwarz criterion	4.780199
Log likelihood	-260.5929	Hannan-Quinn criter.	4.668009
F-statistic	5.672136	Durbin-Watson stat	1.805868
Prob(F-statistic)	0.000013		

**Significant at %10 level, **Significant at %5 level.*

To understand the regression results, we need to consider each variable in relation to soccer as well as their significances mathematically. The variable *bdg* (budget) seems to be highly insignificance in *Table 4.* The budget of a club surprisingly does not matter[17]. Most successful clubs in the world have high budgets and therefore this result is unanticipated. One reason might be that the budget is captured by the new coach ability, average team quality and other variables that are correlated to the budget of a club. A club that has a high budget might perform better and thus may already have performed better which can be seen in the average team quality variable. In addition, the higher the coach ability the more costly the coach is likely to be and this might also include the

[17] The coefficient of the budget variable is infinitesimal. At first glance, it might seem odd, however the budget is in millions of dollars and the dependent variable (*xp)* has a numerical range from 0 to 12.

budget effect that we were expecting. In order to evaluate this concept, a quick scatter plot was constructed as can be seen in *Figure 6* (See appendix for other related figures). Although a clear trend is not visible in the illustration of the relationship between these two variables, one can unambiguously see that teams with higher budgets do hire the most skillful coaches. Moreover, a similar pattern can be observed between *budget* and *team average quality*. Therefore, this variable is also excluded from the model, thereby arriving at our edited version of the short-run and long-run models.

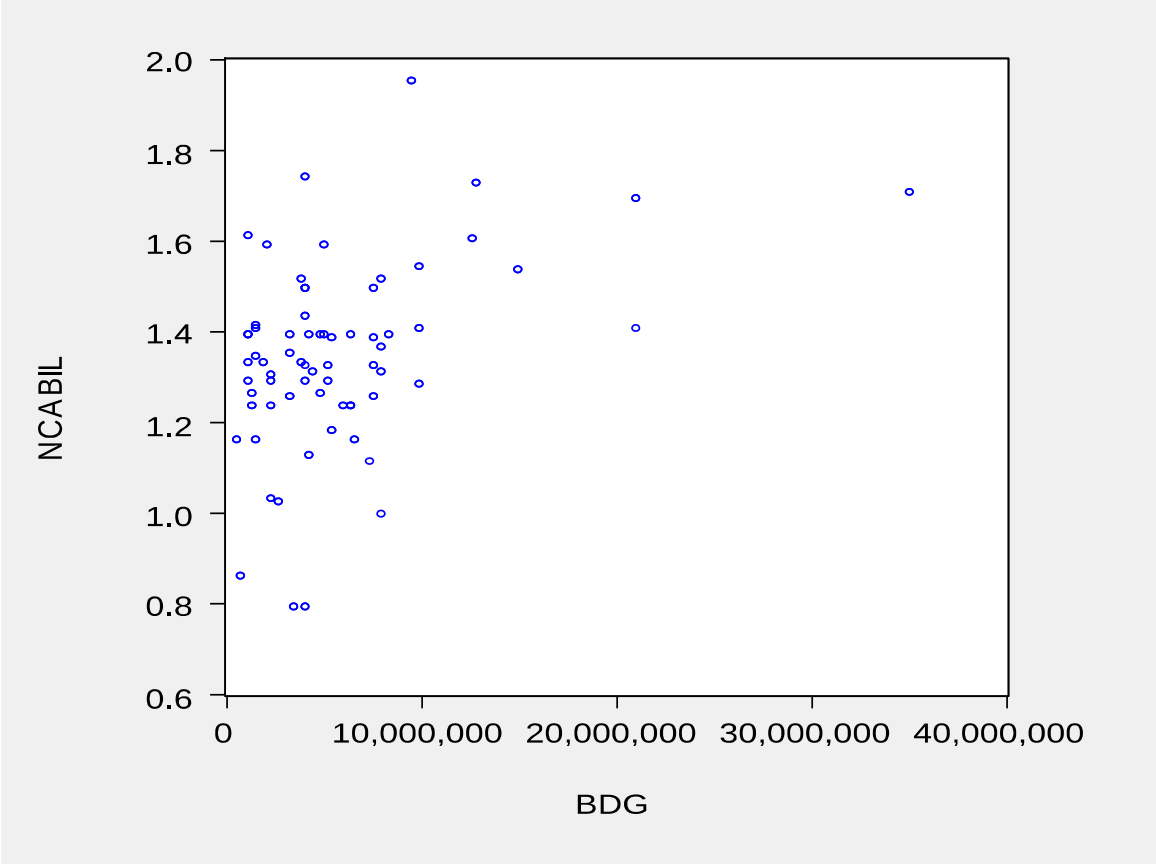


Figure 6. Budget vs. New Coach Ability

Coefficient Analysis & Structural Differences

Some of the observations in the model include teams that are just promoted that year into the Turkish 1st Division. Promoted teams tend to appear easy as targets in general

(Tena, 2006). This is due to their inexperience in the first division against strong and historical teams that have much more qualified players. This inexperience, at times, causes fear and critical mistakes in games. Therefore, there might be a structural difference between the functional form of promoted teams and that of all the other teams. The dummy variable *prom* was multiplied with *tavq* and *ncabil* to see whether promoted teams have different weights for their past performance and their newly hired coach. It might be that a highly successful coach might not be able to reflect his skills on the pitch by coaching an inexperienced team. Furthermore, the average quality of a promoted team might not reflect the true quality of the team if the team has only played a few seasons in the Turkish First Division. To check if there is a structural difference, that is a different slope and different intercept for promoted teams. The Wald test was used to check the joint significance of these variables.

Table 7. Joint Significance – Promoted Teams (Short-Run)

Wald Test:
Null Hypothesis: C(PTQ)=0, C(PNCA)=0, C(PROM)=0
Equation: SHORTRUN

Test Statistic	Value	df	Probability
F-statistic	1.240855	(3, 107)	0.2986
Chi-square	3.722564	3	0.2930

Our claim that promoted teams might be structurally different fails at the five percent significance level. Therefore, we cannot conclude that average historical performance and coach ability have different effects on newly promoted teams nor that they have different intercepts. However, a more interesting result can be seen in the long run model. One might expect that the joint significance of three insignificant variables might be also insignificant but this is not the case in this particular situation.

Table 8. Joint Significance – Promoted Teams (Long-Run)

Wald Test:
Null Hypothesis: C(PTQ)=0, C(PNCA)=0, C(PROM)=0
Equation: LONGRUN

Test Statistic	Value	df	Probability
F-statistic	3.136894	(3, 108)	0.0284**
Chi-square	9.410681	3	0.0243**

***Significant at %5 level.*

In the long-run model, we can conclude that overall promoted teams are in fact different than non-promoted teams. Even though we do not have the ability to individually compare new coach effects, team history effects and/or intercepts in promoted and non-promoted teams, we can definitively make an overall judgment about these particular teams.

Table 9. Long Run Edited Model
Dependent Variable: XPP
Method: Least Squares
Date: 03/13/13 Time: 13:28
Sample: 1 86 88 117
Included observations: 116
Weighting series: GP^2
Weight type: Inverse variance (average scaling)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.314142	0.091852	3.420069	0.0009
NCABIL	0.147542	0.047908	3.079720	0.0026**
EXPRNC	0.000589	0.001248	0.472199	0.6377
TPOS	-0.002491	0.002670	-0.932986	0.3529
TAVQ	-0.003586	0.003340	-1.073631	0.2854
PTQ	-3.20E-05	0.008564	-0.003734	0.9970
PNCA	-0.157042	0.120614	-1.302016	0.1957
PROM	0.152890	0.204345	0.748196	0.4560

Weighted Statistics

R-squared	0.270814	Mean dependent var	0.379419
Adjusted R-squared	0.223552	S.D. dependent var	0.234155
S.E. of regression	0.088508	Akaike info criterion	-1.944978
Sum squared resid	0.846034	Schwarz criterion	-1.755075
Log likelihood	120.8087	Hannan-Quinn criter.	-1.867888
F-statistic	5.730051	Durbin-Watson stat	1.743008
Prob(F-statistic)	0.000012	Weighted mean dep.	0.433950

Unweighted Statistics

R-squared	0.252189	Mean dependent var	0.410333
Adjusted R-squared	0.203720	S.D. dependent var	0.159600
S.E. of regression	0.142419	Sum squared resid	2.190569
Durbin-Watson stat	1.690288		

***Significant at %5 level.*

D. Discussion

After an in-depth econometric analysis, the paper will continue by discussing these results from a soccer fan's point of view. By gathering approaches and ideas from literature in the topic, concrete results were established in both the short-run and long-run models.

The short-run model estimated that among many factors, playing at home is one of the most important factors that increase the performance of a team once a new coach takes over. This result was also observed in Tena & Forrest's (2007) research conducted in the Spanish La Liga[18]. Tena & Forrest (2007) observed a small improvement in games that were played at home after a turnover. In our analysis, we observed *phome* to have a strong positive and highly significant coefficient, which implies that there is indeed *home advantage* in that teams who play more games at home within the four games subsequent to the dismissal, are more likely to perform better than teams who play away. Fan support throughout soccer history has proven to be a strong advantage and our analysis does nothing but support this argument. Our analysis suggests that teams who play one more game at home in the given four games after a dismissal will typically earn one more point.

Another variable that seemed to be highly crucial is the ability of the new coach. This result is intuitive and not surprising; however, what is surprising is the way it is

[18] Spanish 1st Division

measured. Many critics in the field have either not found efficient proxies for coach ability or have used proxies like experience or number of top teams managed (De Paola & Scoppa 2008; Frick & Simmons, 2008). Using average number of points in a coach's entire career has proven to be an efficient and intuitively accurate factor that determines the ex-post performance. New coach ability not only explains performance but also includes many additional factors such as the budget of a team, and the expectations of the team[19]. We see that *ncabil* is significant throughout our analysis similarly to *phome*. Findings suggest that a coach that has one more point per game in his entire career will on average earn around 3.5 points more in a four-game period when he takes over a team.

The last variable that attracts considerable attention is how differently promoted teams' histories affect their performances comparatively to non-promoted teams. Recall that quality is measured by rank so a greater number implies a lower quality. It appears that team average quality in general has a positive coefficient meaning that the more successful the team has performed in the past five years, the worse they will perform in the direct four games after a turnover but the coefficient is not significant. For promoted teams however, this variable has a negative and significant coefficient. The past five-year performance of promoted teams explains considerably more than teams who have been in the division a year ago. The negative coefficient suggests that teams who have performed one ranking higher in earlier seasons will earn on average about 0.41 more points than teams ranked one ranking lower.

In the long run, things get less conclusive as the edited model in the long run yields distinctive results. The only variable that seems to be significant at the five percent level

[19] A team with a higher budget or successful history will hire more skilled coaches in order to be successful.

is the new coach ability. Specifically, teams who hire coaches that average one point higher per game in their careers end up winning 15% more of their games in the remaining of the season. This result surely supports the argument that the new coach ability is the key factor overall in determining long-run performance. Additionally, other variables seem to explain the ex-post performance and are jointly significant but individually it is difficult to measure their marginal influences since they are insignificant individually. Similarly to the short-run model, these variables include experience, team quality, team position, promoted team quality and the promoted dummy.

E. Robustness

In order to effectively claim a conclusion about the factors that ex-post effect, we will test for robustness by answering the same question from a different standpoint. After all, these results could have just fit by coincidence and regression fishing might be a possibility. Hence, we are going to test whether the variables that appeared to be significant in determining ex-post performance will appear to be significant again by regressing different variables that reflect similar concepts on a similar dependent variable. The model is setup through variables that intuitively measure the ex-post performance only, disregarding the ex-ante performance. The setup of the model makes sense though; due to the structure of the model the dependent variables can be adjusted to explain the dependent variable in a proportional fashion. Intuitively, the right-hand side of the model is trying to explain the performance after a turnover, but if we change the left-hand side from ex-post performance to change in performance then we can change some of the independent variables located on the right side of the equation. Hence,

$\beta_{12}X_{tavq} + \beta_{13}X_{tpos}$ can be changed into $\beta_{12}(X_{tavq} - X_{tpos})$, which measures the magnitude of how much the team is under or over performing relatively to the expectations. In addition, the same can be done for $\beta_8 X_{oqante} + \beta_9 X_{oqpost}$, i.e. $\beta_8(X_{oqante} - X_{oqpost})$. The change in performance should be determined by the change in the qualities of the teams played; consequently, this restriction is appropriate for testing in the model. Nonetheless, the difference between opponent qualities was excluded from the model since neither of the variables were in the edited short-run model. The same reasoning can be applied to the old and new coach abilities to get $\beta_8(X_{ncabil} - X_{ocabil})$[20].

Several Wald Tests were run on the short and long run final edited model (with certain restrictions)[21] to see whether the coefficients should be restricted in this way for these pairs of variables. The similar procedure was performed on *tavq* and *tpos.* Looking at history of turnovers and earlier literature, a consensus is that teams are considered to be performing poorly if they are underperforming relative to their expected performance. In 2009-2010 in the Spanish First Division, Real Madrid broke the record of scoring goals in the history of the league and also achieved an all-time most wins at home in their history. But Real Madrid fans were not happy. They have performed well but not better than Barcelona, their biggest rival. Teams in general dismiss managers due to less-than-expected performance and not literally poor performance (De Paola, 2008; Frick & Simmons, 2008; Rowe 2005). Therefore, there is an incentive to try and restrict these

[20] Note that putting the regression in this form forces the ex-ante coefficients to equal the negative of the ex-post coefficients

[21] The edited short run model was taken from the earlier sections and the dependent variable was changed to the difference of performances. Similarly, the edited long run model was also altered.

coefficients. Indeed, the F-statistic that was computed cannot reject the null hypothesis at the five percent significance level. From a soccer-analysis standpoint, this means that these variables could be of same form and their coefficients could have the same magnitude with opposite signs.

We utilize from the same test in order to check whether this same structure is possible with the two variables: *old coach ability* and *new coach ability*.

Table 10. Wald Test – Team Average Quality & Position (SR[22])

Wald Test: C(TPOS)=-C(TAVQ)
Equation: SHORTRUN

Test Statistic	Value	df	Probability
t-statistic	0.031967	109	0.9746
F-statistic	0.001022	(1, 109)	0.9746
Chi-square	0.001022	1	0.9745

The difference between the skills of the coaches should intuitively explain the difference between ex-post and ex-ante performances. The following table depicts the Wald test results for this restriction on the final edited short-run model.

Table 11. Wald Test – New & Old Coach Abilities (SR)

Wald Test: C(NCABIL)=-C(OCABIL)
Equation: SHORTRUN

Test Statistic	Value	df	Probability
t-statistic	0.966890	108	0.3358
F-statistic	0.934876	(1, 108)	0.3358
Chi-square	0.934876	1	0.3336

Indeed, the results are similar to the previous test. That is, the difference version of these variables is suitable for use since we cannot reject the null hypothesis at any reasonable level.

Finally, the last pair of variables that are tested: games played at home after and before a

[22] Short-Run

dismissal (*phome, ahome).* The change of fan support could be proxy for explaining change of performance. Results yield similarly as other tests as can be seen in *Table 12.*

Table 12. Wald Test – Ex-post & Ex-ante Game Locations (SR)
Wald Test:
Equation: SHORTRUN

Test Statistic	Value	df	Probability
t-statistic	0.679756	108	0.4981
F-statistic	0.462069	(1, 108)	0.4981
Chi-square	0.462069	1	0.4967

The regression analysis was run with including the pairs of variables and *Table 13* illustrates the results.

Table *13. Short Run Robustness*
Dependent Variable: XP-XA
Method: Least Squares
Date: 03/13/13 Time: 23:54
Sample: 1 117
Included observations: 117

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.871627	1.307261	1.431716	0.1550
PHOME-AHOME	1.078696	0.421960	2.556392	0.0119**
NCABIL-OCABIL	1.771719	0.901843	1.964553	0.0520*
TAVQ	-0.067048	0.091607	-0.731909	0.4658
PROM*PTQ-TPOS	-0.118388	0.061036	-1.939629	0.0550*
PNCA-OCABIL	0.825980	0.674310	1.224925	0.2232

R-squared	0.138160	Mean dependent var	1.487179
Adjusted R-squared	0.099338	S.D. dependent var	2.993140
S.E. of regression	2.840585	Akaike info criterion	4.975818
Sum squared resid	895.6508	Schwarz criterion	5.117468
Log likelihood	-285.0853	Hannan-Quinn criter.	5.033326
F-statistic	3.558838	Durbin-Watson stat	1.730030
Prob(F-statistic)	0.005056		

***Significant at %5 level. *Significant at %10 level.*

Comparing the results to the edited short-run model, we can see that similar variables are significant. The main three variables that were significant in the short-run model were *phome, ncabil* and *ptq.* Similarly, the significant variables in this regression model with

same variables are *phome-ahome, ncabil-ocabil* and prom**ptq-tpos.* The results support the claim that concepts such as location, coach ability and average team quality (team expectations) are crucial variables when determining performance in the short-run once a turnover occurs.

Similar Wald tests were ran on the long-run model and it turns out that *tavq-tpos, ncabil-ocabil* and *prom*ptq-tpos* can be used for the robustness model (see appendix).

Table 14. Long Run Robustness

Dependent Variable: XPP-XAP
Method: Least Squares
Date: 03/14/13 Time: 00:08
Sample: 1 86 88 117
Included observations: 116
Weighting series: GP^2
Weight type: Inverse variance (average scaling)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.241943	0.049897	-4.848904	0.0000
TAVQ-TPOS	0.002921	0.004388	0.665848	0.5069
NCABIL-OCABIL	0.072287	0.036600	1.975044	0.0508*
EXPRNC	0.004006	0.001670	2.399321	0.5181
PROM*PTQ-TPOS	-0.027033	0.004165	-6.490966	0.0856*
PNCA	-0.142412	0.173888	-0.818987	0.4146
PROM	0.497610	0.248925	1.999033	0.0481**

Weighted Statistics

R-squared	0.426717	Mean dependent var	0.101514
Adjusted R-squared	0.395160	S.D. dependent var	0.181012
S.E. of regression	0.123549	Akaike info criterion	-1.285917
Sum squared resid	1.663804	Schwarz criterion	-1.119752
Log likelihood	81.58319	Hannan-Quinn criter.	-1.218464
F-statistic	13.52214	Durbin-Watson stat	2.151184
Prob(F-statistic)	0.000000	Weighted mean dep.	0.133299

Unweighted Statistics

R-squared	0.056958	Mean dependent var	0.081903
Adjusted R-squared	0.005048	S.D. dependent var	0.187117
S.E. of regression	0.186644	Sum squared resid	3.797120
Durbin-Watson stat	1.728434		

***Significant at %5 level. *Significant at %10 level.*

Once again, our robustness analysis regression is consistent with the edited long-run model. Both models having similar results to the earlier analysis gives us a stronger argument in reaching certain conclusions about the factors that affect ex-post

performance.

F. Conclusion

Dismissals occur quite frequently in businesses and in soccer clubs. Unlike in many businesses, contract terminations can be attributed more to coaches (CEO equivalents) in soccer. Our results have yielded interesting theories related to managerial turnovers in the context of soccer. In the literature, we discussed a few main theories that try to explain manager dismissals and their after-effects. The *scapegoating hypothesis* as well as the *coach effect* and the *vicious cycle hypothesis* are three theories that have been the center of attention for many scholars. Our analysis provides support for the *scapegoating hypothesis* but rules out the *vicious cycle* and *coach effect* hypotheses*;* an embarrassing result for teams who have fired managers in the past.

Through econometric models, a wide array of data was analyzed and interpreted. Firstly, the constructed control group shows the *regression to the mean* effect that casted doubt on the effectiveness of manager terminations. Comparing the results to the regression group data, it was found that firing coaches does not help team performance unambiguously. This rules out the *vicious cycle* claim and *the coach effect.* These theories argued that firing a manager worsens the performance of a team while the latter supports the theory that in the short run there is a honeymoon period in which the team does perform better. The control group analysis provided evidence against both of these theories by showing that statistically the control group post performance is the same as the regression group performance if not worse. The *scapegoat hypothesis* asserts that dismissals take place in order to find a person to blame for the performance. Since we see

that the mid-season dismissals do not matter from our results, *scapegoating* seems to be a high possibility in the *Turkish Super Toto Super League.*

It was found that in the short run, the number of games that are played at home, the new coach abilities and historical team performance are the consistent factors that determine the ex-post performance of a team. This supports the hypothesis that coaches do *matter,* but not necessarily that a termination is helpful. In addition, the *home advantage* seems to play a big role in increasing performance after a turnover. This can be attributed to the excitement of the fans for a change and therefore supporting their team strongly in games played in the home stadium of that team.

In the long-run models, we can see similar results. Coach ability and team quality were still the variables that have seemed to be crucial in determining long-run performance. Moreover, we see that promoted teams are structurally different than other teams, and therefore should be analyzed differently. An interesting result that we have discovered is that promoted teams tend to have different effects. Specifically, coaches and past performances of newly promoted, inexperienced teams tend to have considerably distinctive responses to their new coaches, and do not follow their past performances as much as teams who have been in the division all along.

To conclude, soccer manager dismissals within the 14 years that were analyzed tend to be not effective, supporting the scapegoating hypothesis. In addition, factors that affect ex-post performance in the short and long run are mostly based not only on skill and tradition but also on emotion. Teams perform much better if they are in front of their fans; confidence and the passion of the fans are definitely factors that facilitate performance recoveries. The belief that a CEO is the major determinant of performance is

supported in the context of soccer however a turnover was shown to be ineffective in various cases.

G. Appendix

Variables that were not introduced the Variables section are introduced and explained below.

Ex-Post Location (phome): this variable is similar to **Ex-Ante Location** except that it is collected by recording the first four games under the new coach.

Ex-Post Opponent Quality (oqpost): this variable has the same characteristics as the previous one, except it is formed by data from the four games after the dismissal instead of before.

New Coach Ability (ncabil): this variable is similar to the **Old Coach Ability** with the only difference being that it measures the new coach's ability instead of the one fired.

Multiple (multpl): is formed as the product of **Experience** and **New Coach Ability.** It allows the model to give different coefficients to coaches with zero experience and experienced coaches. It is always included with **New Coach Ability**. Realistically, the average points per game of a coach should not matter as much if he has coached very

few games. On the other hand, a coach with years of experience should be able to reflect his full potential in the current year he is coaching.

Board (brd): equals one if there is a change in the management or board of the club, and zero otherwise. Changes in the board are highly crucial in soccer. The vision of a team, the approach a team takes, is highly dependent on the board. A change in board creates an unstable environment to objectively analyze performances and coach dismissals. Therefore, this dummy variable is created in order to take into account these seldom occurrences in the model.

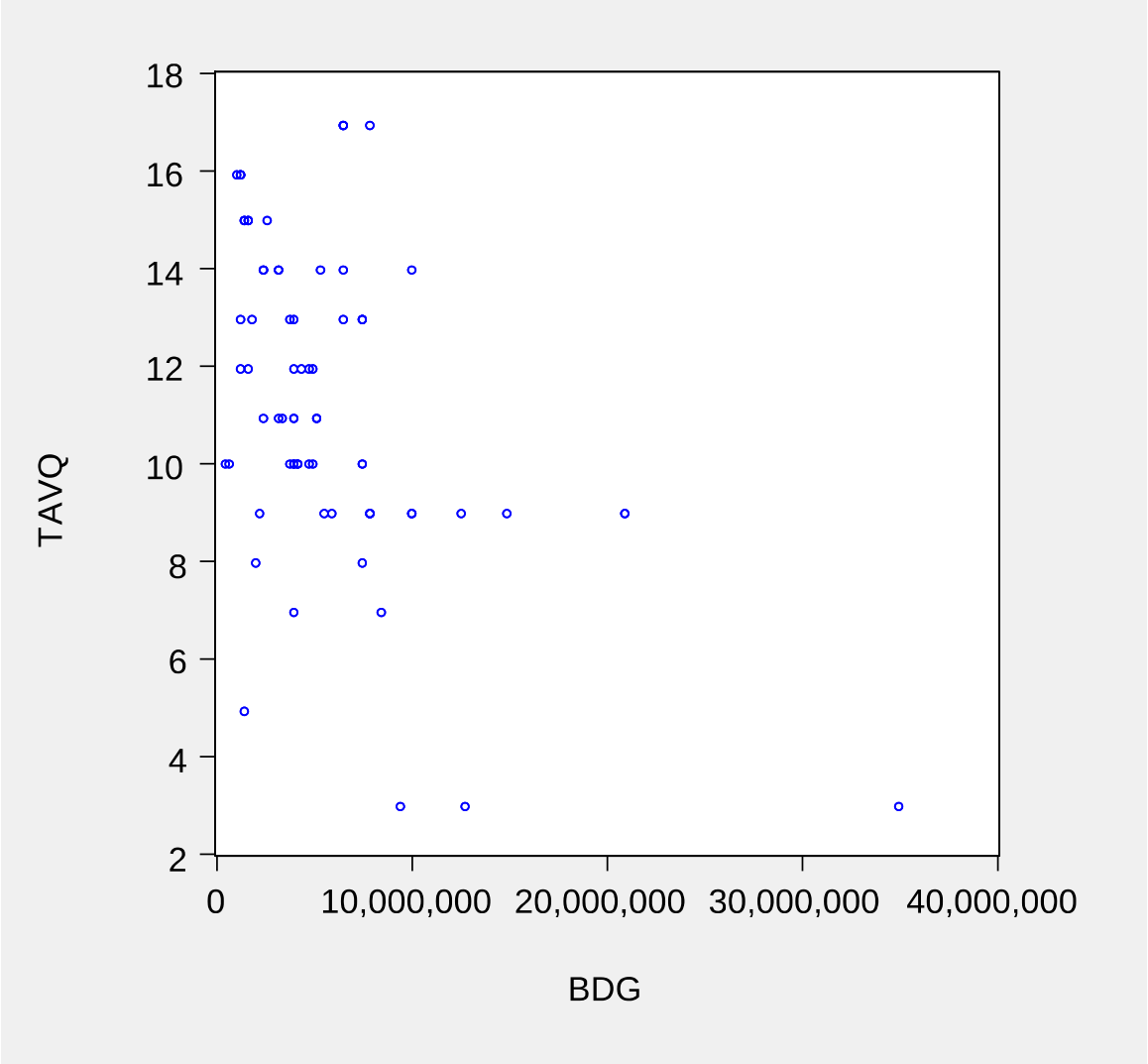
Week (w): is the week in which the manager was fired. A team underperforming in later weeks is more of a problematic situation than a team underperforming in the first few weeks. Therefore, this variable is created to keep track of the timing of the dismissal. Intuitively, a team in, say week 30 (four games until the end of season) might have a brief *coach effect* that would result in a satisfied outcome. The variable is included in all the models in order to take into account of such examples.

Promoted New Coach Ability (pnca): is the product of **Promoted** and **New Coach Ability**. The variable is created to let promoted teams have a different coefficient for their new coaches. It is expected that coaching a newly promoted team is different than coaching a team that has been in the division for a couple of years. Having a stable and developed infrastructure and a confident mentality that is suitable for the first division is highly crucial in order to perform well. Therefore, the model allows us to distinguish the influences of coaches in promoted teams. It is used in both the short run and long run models.

Promoted Average Team Quality (ptq): is the product of **Promoted** and **Average Team Quality**. Most promoted teams have not been in the first division many times; therefore, their past average team qualities might not be an accurate proxy. Thus, this product is introduced to give these promoted teams a different coefficient for their average quality. It will be used in all the versions of the short-run and long-run model.

Figure 6A. Budget vs.Team Average Quality)

Similar to the relationship between coach ability and budget, the relationship between



team average performances is positively correlated.

Table 10A. Wald Test –Team Position & Average Quality (LR[23])

Wald Test:
Equation: LONGRUN

Test Statistic	Value	df	Probability
t-statistic	6.052213	108	0.0227
F-statistic	36.62928	(1, 108)	0.0227
Chi-square	36.62928	1	0.0227

Table 6B. Wald Test – Promoted Team Position & Average Quality (LR)

Wald Test:
Equation: LONGRUN

Test Statistic	Value	df	Probability
t-statistic	2.232566	108	0.1276
F-statistic	4.984352	(1, 108)	0.1276
Chi-square	4.984352	1	0.1256

Table 7A. Wald Test – New & Old Coach Ability (LR)

Wald Test:
Equation: LONGRUN

Test Statistic	Value	df	Probability
t-statistic	1.182911	108	0.2394
F-statistic	1.399278	(1, 108)	0.2394
Chi-square	1.399278	1	0.2368

The long run Wald tests yield similar results. That is, coach abilities and expected and actual performance pairs for promoted and all teams are also included in the model.

Table 8A. Wald Test – Opponent Qualities Ante & Post (SR)

Wald Test:
Equation: SHORTRUN

Test Statistic	Value	df	Probability
t-statistic	-2.214656	107	0.0289
F-statistic	4.904702	(1, 107)	0.0289
Chi-square	4.904702	1	0.0268

[23] Long Run

The only pair of variables that we could not use in the short run robustness analysis was opponent qualities. Indeed the p-value is low enough to reject the null hypothesis.

References

Allen, M. P., Panian, S. K. and Lotz, R. E. (1979). Managerial succession and organizational performance: A recalcitrant problem revisited. *Administrative Science Quarterly*, 24, 167-180.

Audas, R., Dobson, S. and Goddard, J. (1999). Organizational performance and managerial turnover. *Managerial and Decision Economics*, 20, 305-318.

Audas, R., Dobson, S. and Goddard, J. (2002). The impact of managerial change on team performance in professional sports. *Journal of Economics and Business*, 54, 633-650.

Bachan, R., Reilly, B., Witt, R., (2008). The hazard of being an English football league manager: Empirical estimates for three recent league seasons. Journal of the Operational Research Society 59, 884–891.

Balduck AL, Buelens M. (2007). Does sacking the coach help or hinder the team in the short term? Evidence from Belgian soccer. Working Paper., Universiteit Gent.

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

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

Crossan, M. M., Lane, H. W. and White, R. E. (1999). An organizational learning framework: From intuition to institution. *Academy of Management Review*, 24, 522-537.

De Paola, M. and Scoppa V., (2008). "The Effects of Managerial Turnover: Evidence from Coach Dismissals in Italian Soccer Teams," MPRA No. 11030.

Fabianic D. (1994). Managerial change and organizational effectiveness in Major League Baseball: findings for the 80s. *J Motor Beh*; 17: 135-148

Frick, B., Simmons, R., (2008). The impact of managerial quality on organizational performance: Evidence from German soccer. Managerial and Decision Economics 29, 593–600.

Frick, B., Barros, C.P., Passos, J., (2010). Coaching for survival: The hazards of head coach careers in the German ''Bundesliga". Applied Economics, in press.

Gamson, W. A. and Scotch, N. A. (1964). Scapegoating in baseball. *The American Journal of Sociology*, 70, 69-72.

Grusky, O. (1963). Managerial succession and organizational effectiveness. *The American Journal of Sociology*, 69, 21-31.

Greiner, L., T. Cummings and A. Bhambri (2002). 'When new CEOs succeed and fail: 4-D theory of strategic transformation', *Organizational Dynamics*, 32, pp. 1–17.

Hautsch, N., Frick, B., Lehmann, E., Warning, S., (2001). Could I stay or must I go? Team-coach-separations in professional soccer in Germany. Mimeo, Faculty of Management and Economics, Witten/Herdecke University.

Hoffler F, Slimka D., (2003). "Do new brooms sweep clean? When and why dismissing a manager increases the subordinates performances", *European Economic Review*, 47, pp. 877–890.

Hughes, M., Hughes, P., Mellahi, K., Guermat, C., (2010). Short-term versus long-term impact of managers: evidence from the football industry. British Journal of Management 21, 571-589.

Kesner, I.F., Sebora, T.C., (1994). Executive succession: Past, present & future. Journal of Management 20, 327–372.

Khanna, N. and Poulsen, A. (1995). "Managers of financially distressed firms: Villains or scapegoats?" *Journal of Finance*, 50:919-940.

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

Maximiano, S., (2006), "Does replacing a manager improve performance", *mimeo*.

Porter P. K., Scully G. W., (1982). Measuring managerial efficiency: the case of baseball. South EconJ 48:642–650

Rowe W., Cannella A., Rankin D. and Gorman D., (2005), Leader succession and organizational performance: Integrating the common-sense, ritual scapegoating, and vicious-circle succession theories, *The Leadership Quarterly*, 16, pp. 197–219

Salomo S. and Teichmann K. (2000)., The Relationship of Performance and Managerial Succession in the German Premier Soccer League, *European Journal for Sport Management*, 7, pp. 99–119.

"Spain 2009/2010". Rec.Sport.Soccer Statistics Foundation (*RSSSF).* Web. 1 Oct 2012.

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.

"Turkiye Futbol Federasyonu." *Turkish Football Federation.* N.p., Web. 1 Oct. 2012

Wirl, F., Sagmeister S., (2008), "Changing of the guards. New coaches in Austria's premier football league", *Empirica*, pp. 1-12.