Available online at www.sciencedirect.com
SCIENCE DIRECT®
European Journal of Operational Research 148 (2003) 233–246
EUROPEAN JOURNAL OF OPERATIONAL RESEARCH
www.elsevier.com/locate/dsw

Manager to go? Performance dips reconsidered with evidence from Dutch football

Allard Bruinshoofd [a,*], Bas ter Weel [b]

[a] *Department of Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands*
[b] *Maastricht Economic Research Institute on Innovation and Technology, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands*

Abstract

This paper examines whether the forced resignation of managers of Dutch football teams leads to an improvement in the results. We find by analysing 12 years of football in the highest Dutch league that forced resignations are preceded by declines in team performance and followed by improvements in performance. However, the improvement in performance after appointing a new manager does not exceed the seasonal average of both the old and new manager. More importantly, using a control group, it turns out that when the manager would not have been forced to resign, performance would have improved more rapidly. We conclude from this that sacking a manager seems to be neither effective nor efficient in terms of improving team performance.

JEL classification: J24
Keywords: Replacement; Ashenfelter's dip; Evaluation; Difference-in-differences

1. Introduction

Ashenfelter (1978) was among the first researchers estimating the effects of training programs on earnings using reference or control groups. [1] His findings suggested an unpredicted decline in earnings of the different groups of trainees in the months prior to entering the training program. This phenomenon is now known in the literature as "Ashenfelter's dip" or the "pre-program dip" and has been the subject of many empirical studies on the effects of training programs on labor-market performance. The econometric analysis of the impact of these training programs on earnings, however, shows little or no lasting effect for participants in the programs. Recently, Heckman and Smith (1999) have hypothesized that Ashenfelter's dip may be explained by a selection bias. [2] The authors find that methods

* Corresponding author.
E-mail addresses: a.bruinshoofd@algec.unimaas.nl (A. Bruinshoofd), b.terweel@merit.unimaas.nl (B. ter Weel).

[1] The analysis was performed in the United States under the Manpower Development and Training Act (MDTA), using a small sample drawn from the Continuous Work History Sample (CWHS) as a control group.

[2] See also Heckman et al. (1997, 2000). In addition, Heckman and Smith (1999) argue that other variables like employment status are much better in explaining the pre-program dip. They conclude that unemployment dynamics and not earnings or employment dynamics drive participation in training programs.

0377-2217/03/$ - see front matter
doi:10.1016/S0377-2217(02)00680-X

controlling only for earnings dynamics, like the conventional difference-in-differences estimator, do not adequately capture the underlying differences between participants and non-participants. In other words, the assumed exogenous process of assigning agents to a training process is not purely exogenous in the sense that there is a bias in the data due to self-selection or implicit job search among agents. Hence, people suffering a large decline in income prior to qualifying for training might be very eager to enroll in a training program. [3]

The econometric methods have been used to estimate the impact of training programs on earnings and have also been used to estimate and investigate a wide variety of other data and (social) experiments. One interesting and unique way to use the techniques is to investigate how severe the performance dip must be before the board of a firm decides to sack a manager. [4] Particularly in the sports business it is well-known that managers are forced to resign when the results of their teams are below the expected level of performance. The process prior to sacking a manager can potentially be described as a special case of Ashenfelter's dip, because the number of games won or the number of points obtained in the weeks before the manager is laid off is likely to be low and declining on average. Similar to the fact that people experiencing a large decline in income might be eager to enroll in training programs, the board, suffering from disappointing performance by the manager, might be eager to take measures to improve performance by sacking the manager. [5]

During the next few games, after a new manager has replaced the old one, the board expects the team to perform at a higher level than before the change of managers. This "shock effect" might save a team from relegation or might enable qualification for lucrative international tournaments like the Champions League or the UEFA Cup. However, the effectiveness of such hiring and firing policies seems to be unclear. What is often observed is that the results and ranking of a team seem to dramatically increase at first but settle after a while at a level which ex post does not always justify the sacking of the previous manager. This lack of justification can be seen from the fact that the new manager has not significantly increased the average performance of the previous manager. [6] Hence, we might hypothesize that boards often replace managers of poorly performing teams even if the managers cannot be fully blamed for poor performance. [7]

In this paper we investigate a particular and unique "experiment" by investigating hiring and firing policies of Dutch football teams in order to observe whether it was justified that managers were set aside due to poor (short-run) performance. During the period 1988–2000, there have been 125 turnovers in the highest Dutch football league for a diversity of reasons. Taking into consideration that there are 18 teams in the Dutch football league this means that each team had on average seven managers in this period. One important reason for a manager to leave a team,

[3] There have been many extensions and improvements of the methods employed since Ashenfelter's initial approach, e.g. Heckman et al. (1999) for an extensive overview.

[4] There are a few studies investigating whether management turnover leads to improved firm performance. See e.g. Denis and Denis (1995), who examine accounting data for large US firms. They find that in those instances where forced resignations took place, large and significant decreases in operating performance prior to management changes were observed and increases following the appointment of new management. The study, however, does not compare these findings with a control group and is therefore not fully able to investigate what would happen if a sacked manager would have stayed.

[5] The board has also other alternatives. They can decide to hire new and probably better players or change the financial incentives for both players and the manager. However, buying new players is often very expensive and changing the payroll system is not so easy because both players and managers have signed contracts which have to be obeyed. It is therefore that firing the manager and describing him as the "scapegoat" is the most easy way to establish a "shock effect".

[6] Khanna and Poulsen (1995) investigate such cases for firms, which are in financial trouble. These firms appear to sack managers, who cannot be fully blamed for poor performance.

[7] The reason why the board would sack a manager even if he is not solely responsible for the disappointing results, is often to bring into effect a "shock effect". After the sacking the board expects the team to perform better (see also footnote 5).

either voluntarily or compulsory, is because of disappointing performance. The intriguing "experiment" in this paper is the following: the "treatment" group consists of managers, who have been forced to resign because of disappointing results; the "control" group consists of managers, whose position we define as "sackable", but who have remained in control. These are situations in which the performance dynamics are comparable to those after which managers have been forced to resign due to poor performance. Our findings suggest that sacking a manager after poor performance does not lead to an improvement in team performance. Using before–after and difference-in-differences estimators we also find that managers who are allowed to stay in case of a performance dip seem to lead their teams on average more rapidly to successful performance again.

The plan of the paper is as follows. Section 2 shortly discusses the composition of the data and elaborates on the specification of the variables and the construction of the control group. In Section 3 we outline the econometric approach and present our results. Section 4 concludes.

2. Data

In this section we discuss the composition of the data and the construction of the variables we use to analyse the sacking of managers. We first briefly describe the Dutch football league from 1988 to 2000 and the teams included in our analysis. Secondly, we present and discuss the turnover events and pay particular attention to the official motivation of this turnover provided by the clubs. Thirdly, we discuss how to measure performance and how to define performance in such a way that we can apply it to before–after and difference-in-differences estimators. Finally, we carefully define a control group consisting of those situations, which we define as "sackable" but in which the manager has been allowed to stay.

2.1. Dutch football league 1988–2000

The highest Dutch football league—the KPN Eredivisie—consists of 18 teams. The composition of the league changes from year to year because of promotion from and relegation to the second division. Our data consist of teams which played in the highest league during any one season between 1988 and 2000. [8] During a season each team has to play every other team twice (once at home and once away), so that the total number of games for each team during a season is 34. For each win a team receives three points, a draw gives one point, and a loss no points. Hence, the end-of-season team scores lie within a range of 0–102. Every season the team ending 18th relegates to the second division and the winner of the second division is promoted to the KPN Eredivisie. The teams ranking 16th and 17th in the KPN Eredivisie have to play relegation playoff games. [9]

2.2. Forced and voluntarily resignation

There have been many changes in the management of the 18 teams during the years under investigation, i.e. 125 turnovers during the 12 seasons we consider, or an average somewhat above 10 per season. This implies that during the average season, roughly 58% of all teams are faced with manager separation for one reason or another.

The turnover motivations are subdivided into six categories; end of contract (42 cases), poor performance (35), voluntary resignation (25), improved offer of another team (9), disturbed relationship (8), and a group of other motivations (6). Straightforward calculation yields that 3 out of every 10 resignations involves a forced resignation due to poor performance. Hence, during an

[8] These teams are in alphabetical order Ajax, De Graafschap, FC Groningen, FC Twente, FC Utrecht, FC Volendam, Feyenoord, Fortuna Sittard, MVV, NAC, NEC, PSV, RKC Waalwijk, Roda JC, SC Heerenveen, Sparta, Vitesse and Willem II. Bruinshoofd and ter Weel (2001) list these teams according to their performance for each year in the period 1988–2000.

[9] From 1995 on the 3-1-0 points rule is effective. Before 1995 it was 2-1-0. However, for ease of comparison, we use throughout this paper the 3-1-0 rule for all years. The ranking of teams does not change much when we rely upon the 3-1-0 rule instead of relying upon the 2-1-0 rule. Only the total amount of points differs. Also, for the measurement of performance dips the points rule is not important as long as wins yield most points and losses least. See Bruinshoofd and ter Weel (2001) for these tests.

average season, 3 out of 18 teams (17%) have sacked their manager as a direct consequence of disappointing results. This seems to be a rather steep hazard rate for managers, even when compared to the most competitive business environments out there (e.g. Khanna and Poulsen, 1995).

In our analysis, we only focus on forced and voluntary resignations during the season. Since contracts are usually specified to end after a season has finished, this excludes all "end-of-contract" situations. The remaining other situations (improved offer, disturbed relationship, and other) occur too infrequently to be of much relevance in the analysis. For that reason our study focuses on forced and voluntary resignations and is relevant for roughly 50% of total turnover per season and roughly 75% of all resignations that are not due to contract expiry (60 out of 83).

2.3. The measurement of performance

In the league tables, team performance is measured by the number of points obtained during an entire season. During a season, this variable is strictly non-decreasing (excluding extraordinary penalties by the football association). Measured in total number of points earned, performance can increase or stagnate, but never decrease. In illustrating the pre-sack dip, we prefer a performance measure that can decrease when performance stagnates. In addition, the "shock effect" implies that we are looking at a period of time shorter than a full season of football, or accumulation of results to date. [10] An alternative would be to measure performance as points earned per game, so that performance can vary from 0 to 1 and 3 (or low, medium, high or any alternative re-scaling of the payoff to losing, drawing and winning, respectively). The important drawback of this alternative is that performance may get too volatile to be altogether informative and may therefore complicate statistical analysis considerably. Here we opt for some sort of compromise between these two extremist alternatives and define a performance measure that is related to points per game obtained on average during the last four games. This way, we obtain a performance measure that can go down considerably in case of a series of bad games, but at the same time is not too sensitive to an occasional loss (win) in a series of wins (losses).

In order to evaluate whether performance is high or low, we want to make a comparison with some "ordinary" performance level for the relevant team in a particular season. Of course, measuring the ordinary level of performance—which should vary with the composition of the squad, the annual budget, the quality of the manager, player form, etcetera—is a complicated methodological issue. We consider three options here. The most direct measure we can think of proceeds along two steps. First, historical information is used to derive the average number of points obtained per game for teams that finish first, second, third, and so on in the league. Pre-season information on the annual budgets for all teams is then used to compute an expected ranking for the end of the current season and the historical average numbers of points per game are then attached to each ranking. This defines the expected level of this season's performance for a particular team based on its budget. We do not implement this normal performance measure for two reasons. The first reason is that we only have information on the annual budgets for the two most recent seasons and partially for one additional season. Secondly, budgets are likely to have an important impact on performance over the long term, but the role of the annual budget

[10] Koning (2000) takes, besides a short-run analysis, also a longer time horizon. His analysis compares performance in all games (during the season) prior to resignation with performance in all games (during the season) after resignation. He concludes that performance does not increase due to forced resignation. The board of a football team is inclined towards a much shorter time horizon, however. Sometimes as short as the immediate effect in the first game after resignation. In our sample, only 3 out of 10 successors to a sacked manager have won their first game, three more managed a draw. This also suggests that there is no conclusive evidence of an immediate "shock effect" after sacking a manager. In case of voluntary resignation, 4 out of 10 successors managed to win the first game and two managed a draw.

on a season-by-season basis seems to be unclear. [11]

Alternatively, we may use some form of adaptive expectations to generate expectations for the current season, i.e. insert the average number of points per game obtained in the previous season (or an average over a particular number of previous seasons) as the expected performance level for the current season. A clear objection to this approach is the bias that may result from buying and selling players in between seasons, changes in the budget over time, and so on.

The measure of ordinary performance we employ for the remainder of this analysis is the average number of points obtained per game during the current season. In this manner we circumvent the issues of (large) changes in the squad's composition and adjustment of annual budgets in between seasons. However, two potentially important problems have to be born in mind. First, "ordinary" performance is determined in-sample and in the case of a manager resignation, we measure this performance as the average performance of the leaving manager and his successor. [12] Should the resigning manager have performed relatively badly, then obviously his successor has an easy job in "beating" this performance level, so that we may overstate the effect of his resignation. Fortunately, the comparison of performance levels between the resigning manager and his successor seems to be still valid. Secondly, in defining the ordinary performance level as the average for the current season, we implicitly impose the requirement that spells of performance below this average are at some point during the season compensated by off-setting spells of above-average performance. However, neither the timing of the return to ordinary performance nor the spell of above-average performance is a priori restricted to follow directly after the resignation of the manager. For these reasons we consider this measure to be a rather accurate proxy for the performance level at which the potential problems should not interfere with our main interest: the direct performance effect of manager resignation.

As a final step, we compute the relative performance of a particular team at a particular point in time as the four-game point average divided by the seasonal average of points per game. Performance defined in this way provides insight into the performance of a team relative to season average, or "ordinary" performance. [13] Whenever performance exceeds unity, the team performs at a higher level than the ordinary level and vice versa. An advantage of this performance measure is that it has the same interpretation for all teams. [14]

2.4. Construction of the control group

Having defined a measure of performance, we may investigate performance dynamics around manager resignations. In order to evaluate the sense of a resignation, we want to determine its effectiveness and efficiency separately. Effectiveness of a resignation requires that the team emerges better from a resignation in terms of performance. Efficiency of a resignation requires that the effect of the resignation could not have been achieved at lower cost by taking other measures rather than sacking the manager.

In order to evaluate the effectiveness of a resignation, we compare performance both prior to and after resignation. We select a period of four games prior to resignation and four games after resignation as our period of analysis. On the one hand, this seems to be a sufficiently long period to

[11] Consider for illustration the case of Ajax, who have the largest annual budget of all Dutch teams. During the four most recent seasons, Ajax won the league once, but finished behind the top three in the other three seasons. This strongly suggests that the short-term link between annual budgets and performance is far less than perfect.

[12] Measuring the average performance of the old and the new manager separately may generate small sample biases in those instances where resignation is near the start or end of the season.

[13] Performance measured by points per game during the last three and five games does not qualitatively alter the results.

[14] We are aware of another potential problem that this definition may generate: a team may be faced with a bad season. In such cases we are a bit too quick to conclude that a team performs better than expected when in fact what happens is that the team performs persistently less than expected. We do not believe, however, that this potential caveat affects our main results significantly.

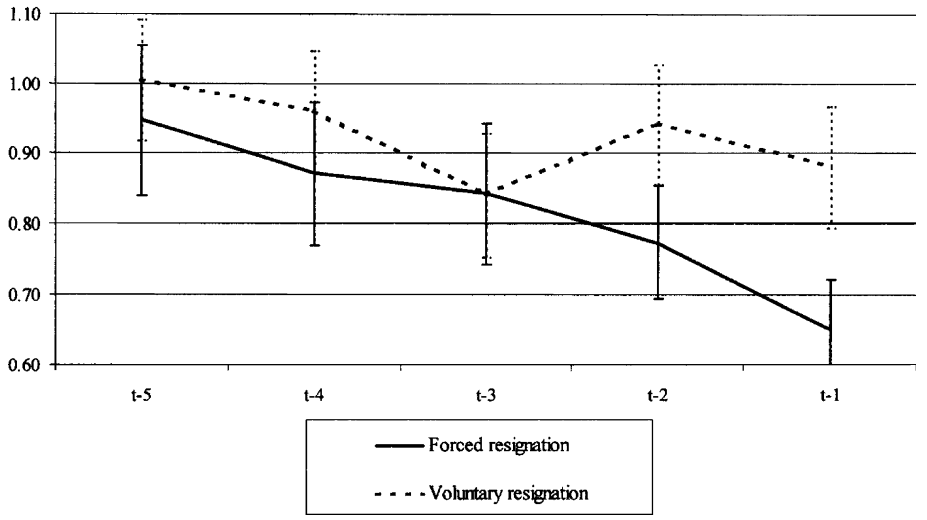

Fig. 1. Performance levels before the date of resignation.

allow for the full effect on our performance measure of a bad streak of four games. On the other hand, it costs us least in terms of lost observations (for resignations close to the start or end of the season). The selected period of analysis requires resignations to take place after game 8 and before game 31 in order to be included in our analysis. [15] What we require in terms of effectiveness is that performance measured 1, 2, 3 or 4 games after resignation exceeds performance 1, 2, 3 or 4 games prior to resignation.

Regarding the assessment of efficiency of resignation, we want to compare it to the most straightforward alternative: do not sack the manager. Obviously, the hypothetical performance of the sacked manager after his forced resignation cannot be observed by definition. Performance after forced resignation when the manager is not sacked is a latent variable, which can only be obtained from a natural experiment. What we can do instead, is construct an indirect measure of these hypothetical performance levels, i.e. we want to know what happens to teams not sacking their managers even though their performance during a four-game period mimics that of a forced resignation. We construct this control group according to the following steps.

First, we present performance levels prior to resignation in Fig. 1. We read from the graph that performance prior to forced resignation drops from 95% to 65% of the season average. For voluntary resignations performance starts out higher (at season average) and drops to close to 90%. For the construction of our control group, we only use the performance information of forced resignations. The major motivation to do so is that this type of resignation is closest related to team performance (the official reading being "sacked due to poor performance"). For voluntary resignations other factors may play a role that do not show up in performance.

Secondly, in terms of performance we observe three distinct features prior to forced resignation: performance is not extremely good to begin with; it declines sharply over a four-game period; and, it ends up at a low level. We apply these criteria to all teams and all seasons to identify those instances in which a four-game period exhibits these three characteristics. In particular,

[15] The reason for this is that we take a four-game average as performance level. This means that games 1, 2, 3 and 4 must be played before we can construct our first performance level. Performance can then decline during the four games 5, 6, 7 and 8 after which the manager may be sacked. In this instance we have exactly sufficient information prior to the resignation. Similarly, at the end of the season, when the manager is sacked in between games 30 and 31, we have games 31, 32, 33 and 34 to examine the potential performance increase after resignation.

Table 1
Number of forced and unforced resignations and performance dips per team: 1988–2000

Club	# Seasons observed	Forced resignation	Voluntary resignation	Performance dip
Ajax	12	2	0	8
De Graafschap	6	0	0	5
FC Groningen	10	4	0	3
FC Twente	12	0	2	8
FC Utrecht	12	3	3	8
FC Volendam	9	0	2	3
Feyenoord	12	4	3	5
Fortuna Sittard	10	1	0	7
MVV	10	0	0	4
NAC	6	1	0	1
NEC	7	2	1	7
PSV	12	0	2	5
RKC Waalwijk	12	3	0	10
Roda JC	12	2	2	4
SC Heerenveen	8	0	0	6
Sparta	12	2	1	4
Vitesse	11	2	2	5
Willem II	12	1	0	10
Total		27	18	103
$\chi^2(17)$ season averages		1.96	1.88	1.71
$\chi^2(17)$ totals		21.67	22.00	18.11

Note: the resignations and performance dips are mutually exclusive in the sense that the figures in the columns concerning forced resignations and voluntary resignations are not included in the column on performance dips. Numbers count the total of observation for the team during the sample period. "$\chi^2(17)$ totals" tests for uniformity of totals across teams, "$\chi^2(17)$ season average" tests for uniformity of dips and resignations per season across teams.

we require an initial performance level of at most 10% above normal; a decline in performance during the following four games of 25% points or more; and, a final performance level of 65% of normal or worse after four games. We make sure that these performance dips (or "sackable" situations) do not overlap with each other or with forced or voluntary resignation. Hence dips and resignations are mutually exclusive in our analysis. [16]

We have a control group of 103 performance dips where the manager is not sacked. In all these cases, performance dynamics during the four games that result in their classification as a dip, validates a forced resignation of the manager when compared to the forced resignation instances. The performance of these teams in the four games after they have "entered" the dip may therefore substitute for the unobservable performance of the resigned manager in the four games after his resignation when he would not have been sacked.

Table 1 summarizes manager resignations and performance dips per team we will use for further analysis. Note that the table does not register all forced and voluntary resignations. As mentioned above, we include only resignations dated after the 8th game and before the 31st game (cf. footnote 15). An important observation resulting from Table 1 and strengthening the representativeness of our control group is that forced resignations, voluntary resignations and performance dips seem to be uniformly distributed across clubs. The relevant $\chi^2(17)$ statistic of the test on uniformity at the bottom of the table never exceeds the 10% critical value (24.8) and for the season-average

[16] We require a distance of four games or more in between any two dips—persistent performance dips would otherwise be weighted too heavily as in some dips the 3 criteria are met for 2 or more periods in a row—and in between any dip and resignation (so that we effectively exclude resignations from the control group).

numbers does not even exceed the 25% critical level (20.5). [17]

3. Before–after analysis and the difference-in-differences estimator

In this section, we examine the mean performance levels of teams around the date of resignation, and consider their implications for before–after estimators of the impact on team performance.

3.1. All resignations

The first three columns in Panel A of Table 2 report the mean scores of sacked managers, managers who have voluntarily left a team and the scores of managers who found themselves in a performance dip but were allowed to stay at the club (the control group). Time is measured in match days prior to and after resignation. [18] From Panel A we observe that performance in case of forced resignation declines from roughly the seasonal average to approximately 65% thereof during the four games leading to forced resignation. The dip reaches its lowest point (is most severe) after the game prior to resignation. In case of a voluntary quit this decline is less pronounced: performance drops from seasonal average to about 90%. Immediately after the manager has involuntarily left the team the performance moves back to slightly below the seasonal average at $t+4$, while a voluntary resignation is followed by a performance increase to approximately 15% above season average. Our control group exhibits a comparable dip in performance around time t; the subsequent upswing, however, is markedly stronger.

In column (4) we first compare performance levels around forced and voluntary resignations. We observe that the performance dip ($t-4$ to $t-1$) in case of forced resignation seems to be somewhat more pronounced but at the same time this difference is not significantly different at error levels higher than 10%. In both instances after the resignation ($t+1$ to $t+4$) performance improves. However, teams sacking a manager seem to keep performing at a lower level than teams who see their manager leave voluntarily. The difference in performance during $t+1$ to $t+4$ remains nearly constant and significantly positive. This indicates that although in both instances performance increases after $t-1$, teams sacking a manager still perform at lower levels relative to teams facing voluntarily resignation. Comparing forced resignations with the control group in column (5) of Table 2, we observe that the performance dip of the control group seems to be more severe. Recovery on the other hand seems to occur much more rapidly. Finally, in column (6) we compare the managers leaving voluntarily with the control group. The most important insight is that the observations around the dip are statistically different from each other while at both ends of the two samples performance is comparable (i.e. the dip is more pronounced, but the initial ($t-5$) and final ($t+4$) performance levels are comparable).

More comprehensive is a comparison of the difference between the pre-sack four-game point average of teams and their hypothetical post-sack recovery, had the manager not been sacked. Following Heckman and Smith (1999), let $Y_{c,u}$ denote the four-game point average without sacking the manager in the period *after* time t ($u > t$) and $Y_{c,-u}$ denote the four-game point average without sacking the manager in the period *before* time t. Similarly, let $Y_{s,u}$ denote the point average with sacking the manager in the period after time t and $Y_{s,-u}$ the point average with sacking the manager in the period before time t. Let $D=1$ for teams who find themselves in a performance dip and $D=0$ otherwise, and let $R=1$ for teams who sack the manager and $R=0$ for teams who do not sack their manager. The experimental impact is then defined as

[17] We might conjecture that "big" teams are harsher towards managers when their higher expectations are not met by team performance, while at the same time performance of "small" teams is more volatile. In this scenario, the resignations are concentrated among the big teams and performance dips among small team, an observation that obviously would cast doubt upon the representativeness of our control group.

[18] Note that time t is non-existent because managers are sacked in between games, not during a game.

Table 2
Team performance around forced and voluntary resignations and performance dips[a]

	Forced	Voluntary	Control group	Forced–vol.	Forced–contr.	Vol.–contr.
Panel A: Performance levels[b]						
$t-5$	0.948*** (0.108)	1.005*** (0.087)	0.921*** (0.014)	−0.057 (0.139)	0.027 (0.109)	0.084 (0.088)
$t-4$	0.871*** (0.102)	0.961*** (0.108)	0.827*** (0.030)	−0.090 (0.149)	0.044 (0.106)	0.134 (0.112)
$t-3$	0.843*** (0.100)	0.841*** (0.115)	0.721*** (0.030)	0.002 (0.152)	0.122 (0.104)	0.120 (0.119)
$t-2$	0.774*** (0.079)	0.942*** (0.104)	0.615*** (0.031)	−0.168 (0.131)	0.159* (0.085)	0.327*** (0.109)
$t-1$	0.650*** (0.071)	0.881*** (0.097)	0.409*** (0.019)	−0.231* (0.120)	0.241*** (0.073)	0.472*** (0.099)
$t+1$	0.664*** (0.069)	0.984*** (0.117)	0.585*** (0.034)	−0.320*** (0.136)	0.079 (0.077)	0.399*** (0.122)
$t+2$	0.704*** (0.084)	1.094*** (0.106)	0.788*** (0.041)	−0.390*** (0.135)	−0.084 (0.093)	0.306*** (0.114)
$t+3$	0.728*** (0.085)	1.069*** (0.106)	0.958*** (0.045)	−0.341*** (0.136)	−0.230** (0.096)	0.111 (0.115)
$t+4$	0.903*** (0.093)	1.176*** (0.078)	1.158*** (0.049)	−0.273*** (0.121)	−0.255** (0.105)	0.018 (0.092)
Panel B: The pre-resignation dip quantified[c]						
Before-1	−0.124* (0.072)	−0.061 (0.074)	−0.206*** (0.028)	−0.063 (0.103)	0.082 (0.077)	0.145* (0.079)
Before-2	−0.193* (0.097)	0.040 (0.122)	−0.303*** (0.030)	−0.233 (0.156)	0.110 (0.102)	0.343*** (0.126)
Before-3	−0.221** (0.098)	−0.080 (0.119)	−0.418*** (0.029)	−0.141 (0.154)	0.197* (0.102)	0.338*** (0.122)
Before-4	−0.298** (0.117)	−0.124 (0.124)	−0.512*** (0.018)	−0.174 (0.170)	0.214* (0.118)	0.388*** (0.125)
After-1	0.014 (0.081)	0.103 (0.067)	0.176*** (0.028)	−0.089 (0.105)	−0.162* (0.086)	−0.073 (0.073)
After-2	0.054 (0.104)	0.213** (0.093)	0.378*** (0.040)	−0.159 (0.140)	−0.324*** (0.111)	−0.165 (0.101)
After-3	0.078 (0.109)	0.188 (0.131)	0.549*** (0.048)	−0.110 (0.170)	−0.471*** (0.119)	−0.361** (0.140)
After-4	0.253* (0.125)	0.295* (0.141)	0.749*** (0.054)	−0.042 (0.188)	−0.496*** (0.136)	−0.454*** (0.151)
Panel C: Before–after analysis and the difference-in-differences estimator[d]						
Before–after-1	0.014 (0.081)	0.103 (0.067)	0.176*** (0.028)	−0.089 (0.105)	−0.162* (0.086)	−0.073 (0.073)
Before–after-2	−0.070 (0.126)	0.152 (0.134)	0.173*** (0.046)	−0.222 (0.184)	−0.243* (0.134)	−0.021 (0.142)
Before–after-3	−0.115 (0.145)	0.228 (0.158)	0.237*** (0.056)	−0.343 (0.214)	−0.352** (0.155)	−0.009 (0.168)
Before–after-4	0.032 (0.139)	0.215 (0.158)	0.331*** (0.059)	−0.183 (0.210)	−0.299** (0.151)	−0.116 (0.169)

* is significant at the 10% level, ** is significant at the 5% level, and *** is significant at the 1% level.

[a] Standard errors in parentheses.

[b] Performance levels are measured relative to season average. Time is measured in games relative to the resignation date, i.e. $t-1$ refers to the last game under the resigning manager. Since managers are not sacked during a game, there is never an observation t.

[c] Before-X refers to the difference in performance levels (Panel A) between $t-X-1$ and $t-1$. After-X refers to the difference in performance levels between $t+X$ and $t-1$. The apparent asymmetry stems from the time-scale used: the performance level at $t-1$ is the same as that at time t, since no games are played in between.

[d] Before–after-X refers to the difference in performance levels between $t+X$ and $t-X$.

$$E(Y_{s,u} - Y_{s,-u}|R = 1, D = 1) - E(Y_{c,u} - Y_{c,-u}|R = 0, D = 1).$$

This is the difference-in-differences in the four-game point average between the "treatment" (sacked managers) and "controls" (managers who are allowed to continue) in period u after the performance dip compared to period $-u$ before the dip.

In Panel B and C of Table 2 we construct the before ($Y_{s,-u}$ and $Y_{c,-u}$ and after ($Y_{s,u}$ and $Y_{c,u}$) estimators separately for the two treatment groups (forced and voluntary resignation) and the control group. From the first three columns it becomes clear that performance before forced resignation is deteriorating significantly, while the performance trend towards a voluntarily quit is also negative but not statistically different from zero. The performance level in the control group is subject to a significant decline prior to the hypothetical sacking. After forced resignation we observe a slight improvement in performance after a while (in particular, only after-4 is significant). Hence, we do not observe an immediate effect of sacking a manager; similar conclusions apply to voluntary quits (significant only for after-2 and after-4). The control group, however, is improving both rapidly and significantly (all After estimates are always positive and significant).

Next, in column (5) of Panel B we compare forced resignations with the controls. We observe that prior to the dip the behaviour of the two groups is statistically similar, but after the dip the performance level of the controls is significantly higher. This suggests that sacking a manager might not be the remedy to solve a crisis. Comparing the controls with voluntarily leaving managers (column (6)) we obtain that the performance prior to leaving is not described by a dip to the same extent as for the control group. After the manager has left, the controls perform much better.

Finally, in Panel C we have constructed the before–after estimates ($Y_{s,u} - Y_{s,-u}$ and $Y_{c,u} - Y_{c,-u}$) and the difference-in-differences estimator ($(Y_{s,u} - Y_{s,-u}) - (Y_{c,u} - Y_{c,-u})$). What is interesting here is that the before–after estimates are only significant for the controls. The before–after estimates are of central importance in evaluating the effectiveness of resignation, because if they turn out insignificant the performance increase is merely the opposite of the dip prior to resignation, i.e. the resignation is not effectively increasing performance. In case of resignation performance indeed turns out to be insignificantly different before and after resignation. This result is consistent with the view that resignations (regardless of being forced or voluntary) are not effective in terms of increasing performance. Of additional interest is the difference-in-differences estimator in column (5). Here we observe that the control group has recovered relatively more rapidly from the performance dip. All coefficients are negative here and seem to indicate (at the 90% level of confidence immediately after the resignation and at the 95% level of confidence after three games) that sacking a manager does not lead to performance improvements compared to allowing him to stay. This finding implies that resignations are also inefficient in boosting performance.

3.2. Successful and unsuccessful resignations

The results presented above suggest that resignations are both inefficient and ineffective as performance improving measures. Though this result holds on average, there are still instances where, ex post, resignation turned out to be very successful. Therefore, we may still be able to discriminate a priori between those resignations that ex post turned out to be successful. In this section we try to gain additional insight into this matter by subdividing the sample into successful and unsuccessful resignations and dips. Successful resignations and dips are defined as those out of which the club emerged better in terms of the before–after-4 measure, i.e. those instances for which performance four games after the resignation exceeded that of four games prior to resignation. Unsuccessful resignations and dips are defined as those for which the before–after-4 estimator is zero or negative.

Panel A in Table 3 presents information on the before–after estimates for the successful forced and unforced resignations and dips. From the first two columns we read that for the 4, 3, 2 and 1 game-

Table 3
Team performance and manager characteristics around successful forced and voluntary resignations and performance dips[a]

	Forced	Voluntary	Control group	Forced–vol.	Forced–contr.	Vol.–contr.
Panel A: Before–after analysis and the difference-in-differences estimator[b]						
Before–after-1	0.207* (0.115)	0.176** (0.064)	0.259*** (0.032)	0.030 (0.132)	−0.052 (0.120)	−0.082 (0.072)
Before–after-2	0.325** (0.137)	0.342** (0.141)	0.330*** (0.056)	−0.017 (0.197)	−0.005 (0.148)	0.012 (0.152)
Before–after-3	0.352** (0.153)	0.497*** (0.155)	0.501*** (0.058)	−0.145 (0.218)	−0.149 (0.163)	−0.004 (0.165)
Before–after-4	0.613*** (0.136)	0.619*** (0.130)	0.658*** (0.053)	−0.006 (0.189)	−0.044 (0.146)	−0.039 (0.141)
Panel B: Manager specific characteristics						
Time employed[c]	32.231*** (5.958)	20.091* (9.660)	54.441*** (6.070)	12.140 (11.350)	−22.210** (8.506)	−34.350*** (11.409)
Total number of dips[d]	2.077*** (0.415)	1.727** (0.727)	2.441*** (0.211)	0.350 (0.838)	−0.364 (0.466)	−0.714 (0.757)
Dips in current season	1.077*** (0.211)	1.000*** (0.270)	1.294*** (0.063)	0.077 (0.342)	−0.217 (0.220)	−0.294 (0.277)

* is significant at the 10% level, ** is significant at the 5% level, and *** is significant at the 1% level.
[a] Successful resignations and performance dips are those out of which the club emerged better in terms of before–after-4, i.e. those situations for which this measure exceeds 0. Standard errors in parentheses.
[b] Before–after-X refers to the difference in performance levels between $t+X$ and $t-X$.
[c] The period the resigning coach was employed by a club, measured in game days.
[d] Total number of performance dips while employed by the club.

periods alike the before–after estimator is positive and in all instances statistically discernibly so at 90% level of confidence or better. From the third column we observe that a similar pattern emerges for the control group. The last two columns show that neither successful forced nor successful voluntary resignations outperform the successful controls. This suggests that sacking—when sacking the manager is ex post successful—results in the same expected performance improvement as when the manager would have been allowed to stay. Hence even ex post successful forced resignations seem to be inefficient.

Panel B of Table 3 reports (i) the number of games the manager was employed by the club when he resigned or experienced a dip, (ii) the total number of dips experienced while employed at this club and (iii) the number of dips experienced during the season in which he is released or faces his most recent dip. The last three columns show that the managers in the control group have been with the club for a significantly longer period of time and they experienced slightly more dips both during their entire careers with the respective club and during the season in which the dip occurs. Hence, although they appear to be slightly poorer managers in terms of dips experienced, they seem to recover from any given dip just as rapidly as would result from a forced or voluntary resignation.

Table 4 presents the same information for unsuccessful resignations and performance dips. Panel A demonstrates that the before–after estimates are significantly negative in most cases for forced resignations and the controls, but insignificant (although negative on average) for voluntary resignations (columns 1–3). Column (5) suggests that unsuccessful forced resignations are structurally more unsuccessful than the controls, while the last column shows that the before–after estimates of voluntary resignations and the controls are not statistically different from each other. Hence unsuccessful forced resignations are inefficient as well and compared to the controls they actually appear counter-productive.

Regarding the manager specific characteristics in Panel B, we focus only on the comparison between forced resignations and the controls (column (5)). Here we see that unsuccessful forced resignations

Table 4
Team performance and manager characteristics around unsuccessful forced and voluntary resignations and performance dips[a]

	Forced	Voluntary	Control group	Forced–vol.	Forced–contr.	Vol.–contr.
Panel A: Before–after analysis and the difference-in-differences estimator[b]						
Before–after-1	−0.165 (0.094)	−0.011 (0.137)	0.014 (0.044)	−0.154 (0.166)	−0.180* (0.104)	−0.026 (0.144)
Before–after-2	−0.436** (0.154)	−0.146 (0.231)	−0.135** (0.052)	−0.290 (0.277)	−0.301* (0.162)	−0.011 (0.237)
Before–after-3	−0.549*** (0.176)	−0.195 (0.266)	−0.248*** (0.063)	−0.354 (0.319)	−0.301 (0.187)	0.054 (0.273)
Before–after-4	−0.507*** (0.112)	−0.421** (0.168)	−0.303*** (0.045)	−0.087 (0.202)	−0.204* (0.120)	−0.117 (0.174)
Panel B: Manager specific characteristics						
Time employed[c]	47.929*** (9.243)	6.714** (2.456)	37.257*** (3.962)	41.214*** (9.564)	10.671 (10.056)	−30.543*** (4.661)
Total number of dips[d]	2.857*** (0.467)	0.143 (0.143)	1.857*** (0.184)	2.714*** (0.488)	1.000 (0.502)	−1.714*** (0.233)
Dips in current season	1.071*** (0.127)	0.143 (0.143)	1.143*** (0.060)	0.929*** (0.191)	−0.071 (0.140)	−1.000*** (0.155)

* is significant at the 10% level, ** is significant at the 5% level, and *** is significant at the 1% level.
[a] Successful resignations and performance dips are those out of which the club emerged better in terms of before–after-4, i.e. those situations for which this measure exceeds 0. Standard errors in parentheses.
[b] Before–after-X refers to the difference in performance levels between $t+X$ and $t-X$.
[c] The period the resigning coach was employed by a club, measured in game days.
[d] Total number of performance dips while employed by the club.

(compared to controls) involve managers who have been with the team somewhat longer, experienced an equal average number of dips during the current season, but during their careers they experienced on average one additional dip. [19] Hence, managers with poorer track records, those we might think are rightfully forced into resignations, actually constitute the resignations that are ex post not only unsuccessful, but even more unsuccessful than the controls. Note that we cannot statistically discriminate ex ante between managers whose forced resignation is ex post successful and those for which it is unsuccessful in terms of manager specific characteristics (compare the first columns of Panels B in Tables 3 and 4). Also note that the share of successful forced resignations in total forced resignations is about 50% while the same share for the controls is roughly 66%.

Our overall reading of the results is the following. When forced resignations are successful, this success does not seem to exceed the success that might have been achieved by not sacking the manager. When forced resignations are unsuccessful, the failure exceeds that which would have to be incurred when the manager would not have been sacked. Since we cannot discriminate between managers whose forced resignation would be a success and those for which it would be a failure, sacking has equal probabilities of success and failure ex ante. The option not to sack during a dip seems to produce a 2/3 probability of success and 1/3 probability of failure. Therefore, in terms of expectations, the expected pay-off of forced resignation seems to be less than that of not forcing the manager into resignation, due to both a higher probability of failure and a larger degree of failure for the forced resignation situation.

4. Concluding remarks

We have investigated performance dynamics around particularly forced resignations during

[19] Voluntary resignations include spells of interim management. This explains both their short average employment spells and the low number of average total (and current season) dips.

twelve seasons in Dutch football. We document that performance improves following both forced and voluntary resignations. However, after close inspection, we find that this effect is likely not to be *caused* by the resignation as our control group exhibits even stronger performance increases. These results are consistent with those obtained in the econometric evaluation of social programs and labour market training programs (e.g. Heckman and Smith, 1999). In the training literature this is often attributed to self-selection by participants even if experiments are seemingly randomized. While it is less conceivable that a manager is inclined to self-select into a forced resignation, our results suggest that the expected effect on average of a forced resignation is a return to the pre-sack performance level. [20] This effect (and more, in fact) is also obtained by not forcing the manager into resignation. Our results shed doubt on the effectiveness and efficiency of forced resignations of which the official reading is "sacked due to poor performance".

Our findings sum up as follows. First, performance deteriorates sharply prior to the resignation and improves rapidly after forced resignation. There is no discernibly positive before–after impact of both forced and voluntary resignations (but there is for the controls). This implies that resignations are ineffective in improving performance. Secondly, difference-in-differences estimates show that the before–after estimates for the control group are statistically discernible and larger than those for forced resignations, but differ little from voluntary resignations. Hence, the data show that forced resignations are inefficient as well as ineffective in improving performance. Thirdly, in cases where forced resignation is ex post effective, it is still inefficient and the same conclusion holds for forced resignations that are ex post ineffective. In particular, where the effect of an ex post successful resignation does not outweigh that of the controls, the failure of ex post ineffective forced resignation *does* outweigh that of the controls. Fourthly, we cannot discriminate a priori between managers whose forced resignation is ex post successful and unsuccessful. These last two findings suggest that in terms of expected valuation, teams do best not to force their managers into resignation in a performance dip.

Overall it turns out that would the manager have been allowed to stay, he would have done slightly better than his successor in improving performance. This is an important result for two additional reasons. First, it seems to become clear there is no such thing as a "shock effect". This underlines that the sacking of a manager seems to be a costly way of signalling there might be something wrong with the team. In addition, it underscores that the manager is often assigned as the scapegoat when performance is temporarily poor (e.g. Khanna and Poulsen, 1995, for business enterprises). Secondly, in large companies CEO's are often blamed for poor performance. In the literature there is little evidence that sacking the CEO leads to improved performance. Hence, an unresolved question is why managers are sacked if it does not materially improve performance. For football our results suggest that it is not his experience (stay at the team) or the ability to deal with performance dips.

[20] Koning (2000) argues that program effects and a series of away games may also substantially influence the performance of a team.

Acknowledgements

We gratefully acknowledge Lex Borghans, Hugo Hollanders, Ruud Koning, Clemens Kool, Bas Straathof and Riccardo Welters for their comments on an earlier draft of this paper. In addition, we want to thank seminar participants at the Johan Cruyff University for their suggestions and discussion.

References

Ashenfelter, O.C., 1978. Estimating the effect of training programs on earnings. Review of Economics and Statistics 60, 47–57.

Bruinshoofd, A., ter Weel, B., 2001. Manager to go? Performance dips reconsidered with evidence from Dutch football. MERIT Research Memorandum 2001-019.

Denis, D.J., Denis, D.K., 1995. Performance changes following top management dismissals. Journal of Finance 50, 1029–1057.

Heckman, J.J., Smith, J.A., 1999. The pre-program earnings dip and the determinants of participation in a social program: Implications for simple program evaluation strategies. Economic Journal 109, 313–348.

Heckman, J.J., Smith, J.A., Clements, N., 1997. Making the most out of programme evaluations and social experiments: Accounting for heterogeneity in programme impacts. Review of Economic Studies 64, 487–535.

Heckman, J.J., LaLonde, R.J., Smith, J.A., 1999. The economics and econometrics of training programs. In: Ashenfelter, O.C., Card, D. (Eds.), Handbook of Labor Economics. North Holland, Amsterdam, pp. 1865–2097.

Heckman, J.J., Hohmann, N., Smith, J.A., with the assistance of Khoo, M., 2000. Substitution and dropout bias in social experiments: A study of an influential social experiment. Quarterly Journal of Economics 115, 651–694.

Khanna, N., Poulsen, A.B., 1995. Managers of financially distressed firms: Villains or scapegoats? Journal of Finance 50, 919–940.

Koning, R., 2000. An econometric evaluation of the firing of a coach on team performance. Mimeo, University of Groningen, September 2000.