UniS

University of Surrey

Discussion Papers in Economics

THE HAZARD OF BEING AN ENGLISH FOOTBALL LEAGUE MANAGER: EMPIRICAL ESTIMATES FROM THE 2002/3 SEASON

By

Ray Bachan
(University of Brighton)

Barry Reilly
(University of Sussex)

&

Robert Witt
(University of Surrey)

DP 16/05

Department of Economics
University of Surrey
Guildford
Surrey GU2 7XH, UK
Telephone +44 (0)1483 689380
Facsimile +44 (0)1483 689548
Web www.econ.surrey.ac.uk
ISSN: 1749-5075

The Hazard of Being an English Football League Manager: Empirical Estimates from the 2002/3 Season

Ray Bachan	Barry Reilly	Robert Witt
Brighton Business School	Department of Economics	Department of Economics
University of Brighton	University of Sussex	University of Surrey
Lewes Road	Falmer	
Brighton	Brighton	Guildford
BN2 4AT	BN1 9QN	GU2 7XH
United Kingdom	United Kingdom	United Kingdom

Abstract

This paper uses data drawn from the English Football League to model hazard rates for club managers in the 2002/3 season. Nearly one-third of managers involuntarily exited employment status with their club in that season. We model the hazard on the basis of a spell at risk, rather than the individual, using a standard logistic model. The role of neglected heterogeneity is also examined using random and fixed effects logistic models within the discrete-time setting. League position at the start of the spell at risk is found to be the most important determinant of a manager's exit. A variety of individual specific human capital covariates were found to be unimportant in determining the hazard and no role for unobservable heterogeneity as captured by random effects was detected.

November, 2005

Keywords: econometrics; sports

Introduction

The high turnover rate among managers in the English Football League is well documented in Audas, Dobson and Goddard (1999). Scully (1994) used a long time series of United States data to examine the relationship between managerial efficiency and managerial survival rates in basketball, (American) football and baseball. In spite of this, there has been a modest literature that attempts to explain the employment survival of football managers. Audas, Dobson and Goddard (*op.cit.*) and Hope (2003) provide rare examples. The former authors use continuous time duration modelling and exploit match-level data over a 25-year period to explain involuntary separations of football managers. On the other hand, Hope (2003) uses Operations Research techniques to determine the optimal time to sack a manager. The purpose of this paper is to add to this modest literature through an examination of managerial exit patterns within a single season. In contrast to Audas, Dobson and Goddard (*op.cit.*), we focus on just one season. The pooling of data over a large number of years potentially imposes a rigid constraint on parameters of interest and may not be entirely sensible given the radical changes the professional game has undergone in England in the last quarter of a century. Our approach emphasizes the role of intra-season performance indicators as determinants of survival.

In contrast to the existing empirical work on this issue, we exploit a discrete-time approach to the hazard modelling. The calendar month within the season provides the measure for the spell at risk. This facilitates use of a standard logistic model and allows an easier introduction of time-varying covariates. In addition, we endeavour to capture neglected heterogeneity through exploiting the panel nature of the spell at risk data.

This paper is structured as follows. The next section contains a description of the econometric methodology used and is followed by a section describing the data. The penultimate section discusses the empirical results and a final section offers some conclusions.

Econometric Methodology

Duration analysis has traditionally been undertaken within a continuous-time framework (see Lancaster (1992)). More recently, however, situating the analysis within a discrete-time context has become popular as it allows for a degree of simplification in the econometric modelling demands. Jenkins (1995), exploiting the work of Allison (1982), suggests a useful way for the estimation of discrete-time duration models using a binary logistic regression model. The approach requires re-organization of the data away from the individual as the unit of observation to the spell at risk of event occurrence. Each individual unit contributes multiple observations to an expanded likelihood function. In the first period an individual unit either stays or exits the state occupied. Those that survive into the second period either remain or exit in that period and so forth into subsequent periods with the numbers surviving, and hence contributing to the likelihood function, diminishing as individuals leave the state in question. The dependent variable in this type of setting could be denoted $y_{i,t}$ and defined as 1 if the i^{th} individual exits the state at time period t, and zero otherwise.

The re-organization of the data allows construction of an unbalanced panel dataset where a maximum of N individuals are observed over T discrete time periods. This could be developed a little further by introducing covariates and separate intercepts for each failure period to capture the baseline hazard. The logistic is a computationally straight-forward transformation that can be used and possesses the added advantage of providing a non-proportional hazard. This may be formulated as:

$$\mathrm{prob}(y_{i,t}=1) = \frac{\exp[\boldsymbol{\beta}'\mathbf{x}_{i,t} + \boldsymbol{\delta}'\mathbf{D}_{i,t}]}{1+\exp[\boldsymbol{\beta}'\mathbf{x}_{i,t} + \boldsymbol{\delta}'\mathbf{D}_{i,t}]} \qquad [1]$$

The dependent variable represents the probability of individual i exiting in an interval around period t conditional on having survived to period t, where $\mathbf{x}_{i,t}$ is a vector of covariates, which may or may not vary over time. The vector $\mathbf{D}_{i,t}$ denotes the baseline hazard, which can be specified by a set of dummy variables for each failure time period (see Sueyoshi (1995)). This provides another advantage to the discrete-time approach in that the estimates for the baseline hazard are derived directly as part of the estimation procedure. This econometric approach has the merit of simplicity and allows an easier introduction of time-varying covariates than parametric continuous-time duration models. This latter advantage has some relevance in the current application given the potential for a strong correlation between the hazard of interest and time-varying performance indicators.

The discrete approach, though computationally straightforward, is not free of criticism. The first relates to the inflation of the sample size through the re-organization of the data. Though a legitimate concern, the estimates obtained remain maximum likelihood and retain the asymptotic properties of such an estimator. The second relates to the treatment of the repeated observations as if they were independent of each other. This is clearly not the case given the data re-organization and correlation across observations potentially introduces some degree of inefficiency in the estimates and a potential downward bias in the sampling variance. However, as Robinson (1982) demonstrates with a probit model, ignoring the correlation across time does not affect the consistency of the maximum likelihood estimates. Following Allison (1982), we exploit an approach within the discrete-time framework that provides a solution to this particular problem (see below).

Neglected heterogeneity, a potentially important issue in duration analysis, could also be addressed by exploiting the panel nature of the data. In particular, the binary logistic model could be re-specified to allow for individual-specific fixed effects as follows:

$$\text{prob}(y_{i,t}=1) = \frac{\exp[\boldsymbol{\beta}'\mathbf{x}_{i,t} + \alpha_i]}{1 + \exp[\boldsymbol{\beta}'\mathbf{x}_{i,t} + \alpha_i]} \qquad [2]$$

The construction of the relevant likelihood for this logistic model and its maximization with respect to the $\boldsymbol{\beta}$ vector and α_i terms yields consistent estimators only if the number of time periods goes to infinity. In empirical applications this is impractical and, with a fixed T, estimates of the fixed effects and the parameter vector are inconsistent. The Monte Carlo experiments reported by Katz (2001) suggest that the scale of the bias may attenuate with T > 15. The inconsistency arises because the number of incidental parameters increases without bound as N $\rightarrow \infty$. (Chamberlain, 1980) suggests maximizing a log-likelihood function conditional on a set of minimal sufficient statistics that help sweep out the fixed effects. A sufficient statistic in the case of the incidental parameters for the logit is $\sum_{t=1}^{T} y_{i,t}$. The model parameters are only identified through the 'within' dimension of the data.

As in the case of the fixed effects linear regression model, only time-varying covariates are permissible with this procedure. The groups of data points used to construct the estimator for $\boldsymbol{\beta}$ are confined to those that exhibit a change in status. The groups that exhibit no change in status are discarded as they provide no information about the vector $\boldsymbol{\beta}$. The use of this procedure requires belief that all the information required for $\boldsymbol{\beta}$ is contained in the non-discarded data. This is a fairly stringent requirement and may not be appropriate for all applications. Another downside to this approach is that the fixed effects cannot be retrieved for further interrogation or analysis. A crude way of overcoming this problem would be to estimate a Linear Probability Model with fixed effects. However, this procedure is inherently heteroscedastic and will have a strong tendency to predict outside the [0,1] boundary given low hazard rate values and this approach is not pursued here.

In the context of the fixed effects model, it is feasible to test for homogeneity across individual effects using a metric version of a Hausman-type test. This entails testing for the difference in estimated coefficients for the covariates in specification [2] compared to those from a pooled logit model with the baseline hazard replaced by a single constant term. It is interpretable as a test for the joint statistical significance of the fixed effects.

The neglected heterogeneity could also be captured through treating the omitted factors as random. This is more within the duration modelling tradition than the fixed effects approach. Assume that the disturbances in the underlying latent relationship are generated by the following process: $e_{i,t} = u_i + v_{i,t}$ where $e_{i,t} \sim N(0, 1)$. This casts the analysis in terms of a probit model but the resultant likelihood function for such an approach is complicated involving T-fold integrals and generally unfeasible with T > 3. A useful device popularised by Butler and Moffitt (1982) reduces the problem to one containing T one-dimensional integrals. This is achieved by integrating u_i out of the joint density function defined as $f(e_{i,1}, e_{i,2,\ldots\ldots}, e_{i,T}, u_i)$. After some tedious manipulation, this gives rise to a random effects probit model with a relatively straight-forward likelihood function (see Greene (2000) for further details). The model is easily adaptable to the logit case using a logistic transformation. The random effects specification could then be written as:

$$\text{prob}(y_{i,t}=1) = \frac{\exp[\boldsymbol{\beta}'\mathbf{x}_{i,t} + u_i]}{1+\exp[\boldsymbol{\beta}'\mathbf{x}_{i,t} + u_i]} \qquad [3]$$

with $u_i \sim N(0, \sigma_u^2)$, where $\sigma_u^2 = \dfrac{\rho}{1-\rho}$ and ρ is the correlation in unobservables across the time units. The u_i term varies across the individuals and is assumed to follow a normal distribution. This could be viewed as more plausible than the Gamma distribution assumed for such heterogeneity in the early duration literature (for example, see Lancaster (1979)). The random effects logistic model can be estimated by maximum likelihood techniques. The additional merit of this model is that it provides a basis for testing the proposition of no cross-period correlations (i.e., ρ=0) and, implicitly, whether σ_u^2 is also zero. It thus provides a formal test for the presence of random effects. If ρ=0 then the estimates should be comparable to estimates from a pooled logistic model.

Data

The empirical analysis uses data collected from the Sky Sports Football Yearbook 2003/4, which includes information on all English League clubs for the 2002/3 season. We use data on 91 of the 92 clubs in the English League for that season. The one omission is Boston Utd., whose manager was banned in late July 2002 and thus does not feature in our analysis. The discrete approach we adopt uses the month as the measure for the spell at risk and there are nine of these in the football season. The choice of time interval is arbitrary but anything shorter did not prove sensible for the construction of the time-varying performance covariates of interest to us. Fortunately, all the managerial exits in this season are interpretable as involuntary separations. Our analysis is only concerned with a single risk and whether the individual exits the state or not. The issue of competing risks is not explored here given data limitations.

We model time to failure starting in early August and concluding in late April as we are only interested in within season separations. This implies that if a manager has survived to the end of April we ignore whether he gets sacked in May or later in the summer close season. The sackings in May are ignored as they might be taken to reflect factors other than within season performance. Thus, we start with the maximum N=91 managers and for those that survive to the end of April, we have a maximum of T=9. Over the period covered 22 out of the initial 91 managers involuntarily exited their employment state with their clubs. The stacked (or panel) data inflate to 736 observations reflecting the loss of individuals due to exits.

The mean of the dependent variable over the entire sample is 0.030. The Kaplan-Meier hazard rates for each of the nine months are reported in table 1 with standard errors in parentheses where relevant. The raw data suggest the peak exit month was October. There were no exits in the months of August, December or February and this curtails how the baseline hazard is specified in the pooled logit model (see below)

The use of the discrete-time approach outlined allows a very simple introduction of time-varying monthly measures in our analysis. A variety of different measures were initially explored to determine their use as covariates in the hazard function. These included an extensive set of time-varying performance measures and time-invariant personal and club-level characteristics (see below). However, only a small sub-set of these covariates exerted any independent influence on the hazard and for brevity, the reported analysis is restricted to those that recorded some effect in one or other of the models estimated. These comprise two time-varying covariates and one time-invariant covariate reported in table 2. The role of other covariates is discussed later.

The position (15^{th} or lower) used to determine whether a team is in the relegation zone or not is broader than conventionally understood. Although this cut-off point is arbitrary it is felt that it adequately captures league under-performance. It is anticipated that this covariate will exert a positive effect on the hazard. The average monthly attendance at home games may be taken to provide a barometer of support for the manager. More importantly, the attendance rate provides the club board the financial incentive to either retain or dispense with the services of the current manager. The greater the attendance rate, the lower is the hazard likely to be. These two measures are based on realizations at the start of the risk month. In the case of the August spell at risk the realisations for the two time-varying variables relate to the league position at the end of the last season and the attendance figure is averaged over the month of April in the previous season. The time-invariant characteristic is a binary measure that captures whether a manager was capped at senior international level during his professional playing career.

Our econometric approach allows for the introduction of factors that capture unobservables. These unobservable effects could be taken to reflect individual heterogeneity that capture coaching ability, man-management skills or other motivational attributes that are assumed to vary across individuals. These omitted factors might also be club-specific in nature and reflect the attitude, patience, indulgence or otherwise of those that have the responsibility for running the club. The omitted factors are separately treated as either fixed or random in the econometric specifications. The importance of these factors is tested empirically.

Empirical Results

The absence of exits in three of the nine months (see table 1) implies that a flexible baseline hazard based on intercepts for individual months does not provide a complete replacement for the constant term in the econometric specification. Therefore, the baseline hazard used in this study conflates the months into pairs for the first six months and allows a final intercept to capture the February to April risk period. This treatment allows for a reasonably flexible baseline hazard that fully replaces the constant term.

The first column of table 3 reports maximum likelihood estimates using the above-defined baseline hazard specification excluding any covariates. The estimates give the effect on the log odds ratio of exit. Once transformed into probability effects these provide Kaplan-Meier estimates but, given the conflation, are not directly comparable to those reported in table 1 above. The relevant transformation is $[1+ \exp[-\beta]]^{-1}$ where β is the logistic coefficient for the period at risk. The highest risk period is October/November where the estimated exit probability is 0.064 – roughly an average of the separate Kaplan-Meier estimates reported in table 1 for these two months.

The introduction of the three covariates (see column two) predictably alters the baseline hazard estimates. However, the October/November period remains the one containing the highest exit risk but the estimated *ceteris paribus* hazard falls by one-half to 0.033. The most important covariate determining the hazard, in terms of both statistical significance and magnitude, is reserved for being in the relegation zone at the start of the risk month. The estimated coefficient suggests that being in this zone raises the probability of exit by 3.6 percentage points – a sizeable effect given the sample mean value for the dependent variable of approximately 0.03 in proportional terms. The estimated effect for the attendance rate, the other time-varying covariate in the specification, is poorly determined with an asymptotic t-value close to unity. The time-invariant covariate capturing whether the current manager was capped in his playing career registers an estimated effect that is just within the boundary of statistical significance at the 10% level using a two-tailed test. The impact effect suggests that, on average and *ceteris paribus*, a manager in this category is 1.4 percentage points more likely to exit the employment state than someone without such a background. This may be explained by club recruitment policy erroneously inferring the existence of managerial and coaching qualities from information on an individual's playing career.

The estimates reported in the third column of table 3 are based on a pooled logit model specification where a constant replaces the baseline hazard. The specifications in columns two and three can now be tested against one another using a conventional likelihood ratio test. The null hypothesis comprises the model reported in the third column of this table. The resultant likelihood ratio test is computed as $11.4 \sim \chi^2_3$, which is statistically significant at the 0.01 level or better. This suggests that the pooled logistic model with the flexible baseline hazard is more congruent with the data than the model containing a constant term.

The estimates for the fixed effects model are reported in column four of table 3. Given the nature of the estimator only time-varying covariates are permissible. The estimator only uses information for those cases where the status changed. This comprises the 22 managers sacked and consists of only 115 observations in total. Once the fixed effects are controlled for, the model estimates suggest a role for attendance but none for league position at the start of the risk month. A Hausman test can be used to determine whether the estimated coefficients for the covariates between the fixed effects model (column four) and a comparable pooled logit model containing the two time-varying covariates (and a constant term) are statistically different from zero. Under the hull hypothesis both estimators are consistent but Chamberlain's fixed effects estimator is inefficient. Given that the log-likelihood function for the fixed effects model is conditional by construction and discards information, conventional likelihood ratio tests cannot be used for this purpose. The null hypothesis comprises homogeneity in the individual-level intercepts. The resultant test value is estimated at $5.644 \sim \chi^2_2$ and the corresponding prob-value is 0.059. Thus, the null is marginally rejected by the data suggesting statistical evidence of variation in the fixed effects across individual managers and, by extension, the joint statistical significance of these effects. However, the fact that in this application a large volume of information is discarded in the construction of the fixed effects estimator provides pause for some interpretational caution here.

The estimates reported in column five of table 3 are based on a random effects logistic model with a constant term. The estimated coefficients for σ_u and hence ρ are both negligible and poorly determined in a statistical sense. This suggests no cross-period correlation and thus, implicitly, no evidence of random effects (or neglected heterogeneity) in the estimated regression model. The estimated effects for the covariates are thus identical to those reported for the pooled logit specification containing a constant term (see column three). It should be noted that replacing the constant term by the flexible baseline hazard in this final specification does not alter the finding in regard to neglected heterogeneity.

The role of a set of other variables on the hazard of exit is now briefly investigated. All but one is time-varying in nature. Given the absence of any evidence of random effects, our preferred model is the pooled logit with the flexible baseline hazard. The variables of interest are introduced separately into this model to assess their statistical significance using either an asymptotic t-test or a likelihood ratio test, and then subject to a joint statistical test overall. The results are reported in table 4. There is no variation in the hazard rates across the division a league club plays. The age of the manager at the start of the season, his length of managerial service at the club, the number of clubs he has managed prior to the current one, and whether or not he is English exert no independent influence on the hazard. In addition, a good FA cup run is unlikely to provide respite to a manager under-performing in the league. And managers who once played for the club they currently manage are not subject to any degree of favouritism in regard to their employment status. The null hypothesis of no joint effects for these variables overall is also upheld by the data.

Conclusions

This short paper used data drawn from the English League to model hazard rates for 91 club managers in the 2002/3 season. Nearly one-third of managers involuntarily exited their employment status in that season. We used a suggestion by Allison (1982) and Jenkins (1995) to model the hazard on the basis of the spell at risk, rather than the individual, using a standard logistic model. We defined the spell of risk in terms of months of the football season. We also examined the role of neglected heterogeneity using both random and fixed effects logistic models in this setting.

The raw data suggest that October was a high risk month for involuntary departures. There was no evidence that controlling for neglected heterogeneity through random effects was important to understanding the determinants of exit. In addition, we are not entirely convinced that the use of Chamberlain's fixed effects estimator to model neglected heterogeneity was all that informative in our application given the volume of information discarded.

Our preferred specification was a pooled logistic model with a flexible baseline hazard. In using this model, we found that the hazard rates were affected by a small set of covariates with the most important a time-varying one defined around success on the field of play. In this regard, our findings resonate with those of Audas, Dobson and Goddard (*op.cit.*). In particular, league position was the most important determinant of an individual manager's exit. The average *ceteris paribus* impact effect on the hazard of being in the league's relegation zone (broadly defined) raised the probability of exit by 3.6 percentage point. It should come as little surprise to English Football League mangers that the key to survival is avoiding a lowly position in the league within which they ply their trade. No independent role for managerial age, experience, length of service or ethnicity was detected. The fact that a manager had a playing relationship with the club in the past also counted for little. Those that run English League football clubs appear understandably to be characterised by a distinct lack of sentiment.

Finally, we believe that the discrete-time approach adopted in this study appears eminently suitable to modelling hazard rates in those applications where fast-changing performance indicators are potentially important determinants of managerial survival. This would generally be the case in most professional sports labour market applications. In addition, the use of random effects within the discrete-time framework to capture neglected heterogeneity provides an added dimension that is relatively easy to implement econometrically. The arbitrary choice of a month to represent the spell at risk could also be relaxed by the interested investigator to allow for a much finer discrete unit based on weeks or even days at risk. However, whether any of the foregoing enhances the empirical analysis is likely to be application dependent.

REFERENCES

Allison PD (1982). Discrete-time Methods for the Analysis of Event Histories, Chapter 2 in *Sociological Methodology 1982*, S.Leinhardt (editor), pp.61-97, Jossey-Bass Publishers, San Francisco.

Audas R, Dobson S, and Goddard J (1999). Organizational Performance and Managerial Turnover, *Managerial and Decision Economics* **20**: 305-318.

Butler JS and Moffitt R (1982). A Computationally Efficient Quadrature Procedure for the One-factor Multinomial Probit Model, *Econometrica* **50**(3): 761-764.

Chamberlain G (1980). Analysis of Covariance with Qualitative Data, *Review of Economic Studies* **XLVII**: 225-238.

Greene WH (2000). *Econometric Analysis*, Fourth Edition, International Edition, Prentice Hall.

Hope C (2003). When Should You Sack a Football Manager? Results from a Simple Model Applied to the English Premiership, *J Opl Res Soc* **54**: 1167-1176.

Jenkins S (1995). Easy Estimation Methods for Discrete-Time Duration Models, *Oxford Bulletin of Economics and Statistics* **57**: 129-138.

Katz E (2001). Bias in Conditional and Unconditional Fixed Effects Logit Estimation, *Political Analysis* **9**(4): 379-384.

Lancaster T (1979). Econometric Methods for the Duration of Unemployment, *Econometrica* **47**: 939-956.

Lancaster T (1992). *The Econometric Analysis of Transition Data*, Cambridge University Press.

Robinson PM (1982). On the Asymptotic Properties of Estimators of Models Containing Limited Dependent Variables, *Econometrica* **50**: 27-41.

Scully GW (1994). Managerial Efficiency and Survivability in Professional Team Sports, *Managerial and Decision Economics* **15**: 403-411.

Sueyoshi GT (1995). A Class of Binary Response Models for Grouped Duration Data, *Journal of Applied Econometrics* **10**: 411-431.

Table 1:Kaplan Meier Hazard Rates

Aug.	Sept.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.
0.000	0.011 (0.018)	0.100 (0.019)	0.025 (0.019)	0.000	0.038 (0.019)	0.000	0.039 (0.019)	0.055 (0.020)

Table 2: The Covariates

Variable Name	Description of Covariates	Mean
Relegation Zone	=1 if the league position at the start of the month was 15th place or below; =0 otherwise.	0.335
Attendance	Attendance as a percentage of a club's ground capacity over the previous month.	62.62 (24.36)
International	=1 if the current manager was capped at senior international level as a player; =0 otherwise	0.276

Notes to table 2:

(a) Numbers reported in parentheses are standard deviations.

(b) The sample averages reported are based on the pooled sample of 736.

Table 3: Maximum Likelihood Estimates of the Hazard Rate

Variables	Pooled Logit with Flexible Baseline Hazard	Pooled Logit with Flexible Baseline Hazard and Covariates	Pooled Logit with Constant and Covariates	Fixed Effects Logit with Covariates	Random Effects Logit with Constant and Covariates
Constant	‡	‡	-3.981* (0.721)	‡	-3.981* (0.724)
August/ September	-5.198* (1.003)	-5.923* (1.248)	‡	‡	‡
October/ November	-2.677* (0.312)	-3.368* (0.762)	‡	‡	‡
December/ January	-3.945* (0.583)	-4.554* (0.917)	‡	‡	‡
February to April	-3.439* (0.384)	-4.032* (0.779)	‡	‡	‡
Relegation Zone	†	1.713* (0.500)	1.671* (0.495)	0.751 (0.671)	1.671* (0.495)
International	†	0.790* (0.477)	0.889* (0.470)	‡	0.889* (0.470)
Attendance	†	-0.009 (0.009)	-0.012 (0.010)	-0.075* (0.043)	-0.012 (0.010)
σ_u	†	†	†	†	0.001 (0.620)
ρ	†	†	†	‡	0.0000 (0.0003)
Log(L)	-93.07	-83.95	-89.63	-30.57	-89.63
Sample Size	736	736	736	736	736

Notes to table 3:

(a) † denotes not used in estimation. (b) ‡ not applicable in estimation. (c) * denotes statistical significance at the 0.10 level or better. (d) Log(L) denotes the log-likelihood value. (e) The log-likelihood value for the fixed effects logit model is not comparable to the other log-likelihood values. (f) The number of discarded observations used to construct the fixed effects estimator was 621.

Table 4: The Set of Unimportant Variables

Variable Description	**Test Statistic**	**Pooled Logit**
Length of Service (in Months) up to July 2002	t-ratio	-0.38
Manager's Age (in Years) at August 2002	t-ratio	-1.16
Manager is English=1; 0 otherwise	t-ratio	0.33
Manager was a player at the current club=1; 0 otherwise	t-ratio	0.70
Number of clubs managed prior to the current club	t-ratio	-0.77
FA Cup Run†	t-ratio	0.94
League Division Dummies	χ^2_3	1.78 (0.62)
Joint Significance of all the above	χ^2_9	3.96 (0.91)

Notes to table 4:
(a) The t-ratios are asymptotic and the relevant critical value at the 0.05 level is ±1.96.
(b) † denotes time-varying covariates.
(c) The FA cup run variable is an ordinal variable that captures the highest round achieved at the start of the risk month.
(d) The pooled logit refers to the specification with the flexible baseline hazard in column two of table 3.
(e) For the chi-squared values the numbers in parentheses refer to the significance level of the test.