Duration Analysis of CEO Turnovers Using Proportional Hazard Model

Yungsan Kim and Keunkwan Ryu*

To analyze CEO turnovers in large U.S. corporations over the period 1981 to 1990, this paper adopts a minimum chi-square estimation of proportional hazard model. A simple specification test of the proportional hazard assumption is also used. Empirical results indicate that elderly CEOs are more likely to be turned over (retirement effect) and worse-than-average CEOs face a lot higher turnover risk (disciplinary effect). Interestingly, performance is found to have non-proportional effects on CEO turnovers across tenure periods. At an earlier tenure as CEO, only good performance matters, increasing the chance of survival. On the other hand, at a later tenure, only bad performance makes a difference, enhancing the possibility of turnovers. (*JEL* Classifications: C41, L20)

I. Introduction

In this paper, we are interested in identifying the termination of CEO (chief executive officer) tenures for disciplinary reasons and in studying its relationship with firm performance. To identify the reasons for CEO turnovers, we consulted the Wall Street Journal Index for each turnover analyzed in this paper. When no explicit reason was available, however, the CEO's age was used as the primary criterion for determining whether the turnover was disciplinary or not. The turnovers which occurred at the age of sixty-four or more were regarded as normal retirements, and others were

* Assistant Professor, Department of Economics, Hanyang University; Assistant Professor, Division of Economics, Seoul National University, respectively. The authors are grateful for helpful comments from Youngjo Lee and an anonymous referee. Keunkwan Ryu acknowledges financial support of the Korea Research Foundation.

[**Seoul Journal of Economics** 1998, Vol. 11, No. 3]

regarded as involuntary (disciplinary).

CEO turnovers are available only up to intervals of length one year, not up to exact points. However, it is more natural to assume that the board of directors' decision on CEO tenure termination can take place at any time. Hence, in the current empirical analysis of CEO turnovers, we face a situation with discrete (interval grouped) data and a continuous model. If covariates are discrete, categorical, or grouped, application of Berkson's *minimum* chi-square estimation yields computationally simple estimators which are asymptotically as efficient as the maximum likelihood estimators.

Ryu (1994) proposes a minimum chi-square estimation of PHM, which is applicable when durations are grouped and covariates are categorical. This paper first introduces his minimum chi-square estimation and specification test, and then applies them to analyze CEO turnovers in large U.S. corporations over the period 1981 to 1990. Specifically, we want to see whether elderly CEOs are more likely to be turned over and whether worse-than-average CEOs face a lot higher turnover risk.

As we expected, CEOs with good performance turn out to stay longer and elder CEOs tend to be more easily turned over. The performance related CEO turnovers are disciplinary in nature, whereas the age related turnovers are more likely to be normal retirement. We also find an interesting asymmetry regarding the effect of performance on CEO turnovers: at an earlier stage of CEO tenure period, only good performance matters (increasing chance of survival), while at a later tenure period, only bad performance makes a difference (increasing chance of turnovers).

The rest of the paper is organized as follows. Section II sets up an estimation framework. Section III explains Ryu's (1994) estimation and test procedures. Empirical analysis of CEO turnovers are carried out in Section IV. Finally Section V concludes the paper.

II. Estimation Framework

For estimation, we will use the minimum chi-square estimation method of the *proportional hazard model* (PHM) (Cox 1972, 1975; Cox and Oakes 1984; Kalbfleisch and Prentice 1980). Under PHM, the hazard rate of a duration is specified as a product of two separate terms: the *baseline hazard* describing the overall shape of the hazard rate over time, and the *proportionality factor* capturing the covariate (regression) effects on hazard rates across different individuals.

Let $T \in R^+$ represent a duration variable of interest. Let $h(t, x) = h_0(t)\exp(x\beta)$ be the hazard rate of duration T, where x is a $1 \times k$ vector of covariates, and β a $k \times 1$ vector of regression coefficients. Quite often, a discrete observation scheme can be represented as an equi-spaced partition Q of the support R^+:$Q = \{0, l, 2l, \cdots, rl, \infty\}$. For expositional simplicity, let us take $l=1$ and $r=2$. The resulting observation scheme is $Q = \{0, 1, 2, \infty\}$. Under Q, durations are available only up to intervals $I_1 = [0, 1)$, $I_2 = [1, 2)$, and $I_3 = [2, \infty)$.

Let α_1 be the probability that duration T survives interval I_1, and let α_2 be the conditional probability that T survives I_2 conditional on that it has already survived I_1. Then by the above proportional hazard assumption, we have

$$\alpha_1 = \exp\left(-\int_0^1 h(t, x)dt\right) = \exp\{-\exp(x\beta + \gamma_1)\}, \tag{1}$$

and

$$\alpha_2 = \exp\left(-\int_1^2 h(t, x)dt\right) = \exp\{-\exp(x\beta + \gamma_2)\}, \tag{2}$$

where

$$\gamma_1 = \log\left(\int_0^1 h_0(t)dt\right),$$

$$\gamma_2 = \log\left(\int_1^2 h_0(t)dt\right).$$

These formulas are originally available in Prentice and Gloeckler (1978).

A grouped duration can be considered as a sequence of binary variables indicating an individual's survival status over a sequence of intervals. By constructing a synthetic data set treating each combination (individual, interval) as a new unit of indexing, we can reduce a grouped duration analysis to a binary choice analysis (Kiefer 1988; Prentice and Gloeckler 1978; Sueyoshi 1991). For each combination (interval, individual), a survivor of the jth ($j=1, 2$) interval independently receives the probability $\alpha_j = F(x\beta + \gamma_j)$ if he or she has covariate vector x, $1 - \alpha_j$ otherwise, where $F(\cdot)$ is defined by $F(y) = \exp\{-\exp(y)\}$.

Let i index each different observation, $i=1, \cdots, n$. Assume that observations are independent. Accordingly, define T_i as the ith duration variable, x_i as the $1\times k$ covariate vector of individual i, α_{1i} and α_{2i} as α_1, α_2, respectively, evaluated at $x=x_i$, $d_{1i}=1$ if T_i survives I_1, $d_{1i}=0$ otherwise, and $d_{2i}=1$ if T_i survives I_2 conditional on $d_{1i}=1$, $d_{2i}=0$ otherwise. Under the observation scheme Q, the ith observation can be summarized as a triple (x_i, d_{1i}, d_{2i}), where d_{1i} and d_{2i} are two *binary dependent variables* indicating whether the individual survives the first and the second interval, respectively.

The log-likelihood function is

$$\sum_{i=1}^{n} [d_{1i}\log\alpha_{1i}+(1-d_{1i})\log(1-\alpha_{1i})+d_{1i}\{d_{2i}\log\alpha_{2i}+(1-d_{2i})\log(1-\alpha_{2i})\}]. \quad (3)$$

This function takes a form of the binary choice analysis, with α_{ji}'s as in the Gompit analysis (Zellner and Lee 1965). By maximizing the log-likelihood function, we can obtain a set of consistent and efficient maximum likelihood estimators. In the next section, when covariates x are categorical, we introduce an alternative estimator which is computationally much simpler, yet asymptotically equivalent to the maximum likelihood estimators.

III. Estimation and Test

To make this paper self-contained, we borrow most of the results in this section from Ryu (1994). For details, readers are encouraged to refer to the original article.

A. Minimum Chi-square Estimation

Minimum chi-square estimator can be defined only when there are many observations for each value of the covariate vector. Often it is described as many observations per cell. A cell is defined as a distinct vector value of covariates. This situation will naturally occur if the nature of covariates is either categorical or somehow aggregated. Note that covariates are often aggregated to save space or to keep anonymity of the respondents.

Suppose that the covariate vector x_i takes on g distinct vector values $x_{(1)}, \cdots, x_{(g)}$. Let n_j, n_{1j}, and n_{2j} be the number of individuals who have covariates $x=x_{(j)}$, the number of individuals who have covariates $x=x_{(j)}$ *and* have survived the first interval I_1, the number of individuals who have covariates $x = x_{(j)}$ *and* have survived *both* intervals I_1 and I_2, respectively. Obviously, we have $n_j \geq n_{1j} \geq n_{2j}$. For each group j, we can estimate each interval survival probability through the relative frequency of those who have survived the corresponding interval. Let $\hat{\alpha}_{1j}$ and $\hat{\alpha}_{2j}$ be those estimators: $\hat{\alpha}_{1j} = n_{1j}/n_j$, and let $\hat{\alpha}_{2j} = n_{2j}/n_{1j}$ (assuming $n_{1j} \neq 0$). Note that $\{\hat{\alpha}_{1j}, \hat{\alpha}_{2j}\}_{j=1}^{g}$ constitute a set of sufficient statistics for the model. In the following discussion, we shall write $x_{(j)}$ as x_j for notational convenience.

By taking the inverse of those interval survival probabilities in (1) and (2), we obtain

$$F^{-1}(\alpha_{1j}) = x_j\beta + \gamma_1, \tag{4}$$

and

$$F^{-1}(\alpha_{2j}) = x_j\beta + \gamma_2, \tag{5}$$

where $F^{-1}(\alpha) = \log\{-\log(\alpha)\}$, the inverse function of $F(x) = \exp\{-\exp(x)\}$.

By expanding $F^{-1}(\hat{\alpha}_{1j}) = \log\{-\log(\hat{\alpha}_{1j})\}$ and $F^{-1}(\hat{\alpha}_{2j}) = \log\{-\log(\hat{\alpha}_{2j})\}$ in a Taylor series around the true α_{1j} and α_{2j}, we obtain

$$\log\{-\log(\hat{\alpha}_{1j})\} = x_j\beta + \gamma_1 + u_{1j}, \tag{6}$$

and

$$\log\{-\log(\hat{\alpha}_{2j})\} = x_j\beta + \gamma_2 + u_{2j}, \tag{7}$$

where u_{1j} and u_{2j} are residual terms in the expansion

$$u_{1j} = \frac{\partial F^{-1}(\alpha_{1j})}{\partial \alpha_{1j}} \bigg|_{\alpha_{1j} = \alpha_1{}^*{}_j} (\hat{\alpha}_{1j} - \alpha_{1j}) = \frac{1}{\alpha_1{}^*{}_j \log(\alpha_1{}^*{}_j)} (\hat{\alpha}_{1j} - \alpha_{1j}),$$

and

$$u_{2j} = \frac{\partial F^{-1}(\alpha_{2j})}{\partial \alpha_{2j}} \bigg|_{\alpha_{2j} = \alpha_2{}^*{}_j} (\hat{\alpha}_{2j} - \alpha_{2j}) = \frac{1}{\alpha_2{}^*{}_j \log(\alpha_2{}^*{}_j)} (\hat{\alpha}_{2j} - \alpha_{2j}),$$

with $\alpha_1{}^*{}_j$ and $\alpha_2{}^*{}_j$ lying between $\hat{\alpha}_{1j}$ and α_{1j}, between $\hat{\alpha}_{2j}$ and α_{2j}, respectively.

The $\hat{\alpha}_{1j}$'s are uncorrelated across j since they are computed from different sets of observations. The first interval survival probability $\hat{\alpha}_{1j}$ has variance $\alpha_{1j}(1 - \alpha_{1j})/n_j$. The $\hat{\alpha}_{2j}$'s are also uncorrelated across j. And $\hat{\alpha}_{2j}$ has variance $\alpha_{2j}(1 - \alpha_{2j})/n_{1j}$, conditional on n_{1j}. These variances can be consistently estimated by replacing the true unknown quantities α_{1j} and α_{2j} by their corresponding estimates, $\hat{\alpha}_{1j}$ and $\hat{\alpha}_{2j}$. Further, we can show that the $\hat{\alpha}_{1j}$ and the $\hat{\alpha}_{2j}$ are uncorrelated.

Therefore, the above two equations (6) and (7) comprise a system of two uncorrelated regression equations with seemingly obvious cross-equation parametric restrictions. Note that the same regression coefficient β appears in both equations and that the error terms are purely heteroscedastic. This is opposite to the case of the so called seemingly unrelated regression equation system (see Zellner 1962) where there are no explicit cross-equation restrictions other than the implicit cross-equation error correlations. The minimum chi-square estimator of $(\beta, \gamma_1, \gamma_2)$ is the weighted least squares estimator applied simultaneously to the above two equations. This estimator has the same asymptotic distribution as the maximum likelihood estimator (see Lee 1992).

Using matrix notation, these two equations can be written as

$$y_1 = Z\theta_1 + u_1, \tag{8}$$

and

$$y_2 = Z\theta_2 + u_2, \tag{9}$$

where $y_1 = (\log\{-\log(\hat{\alpha}_{11})\}, \cdots, \log\{-\log(\hat{\alpha}_{1g})\})'$, $\theta_1 = (\beta' : \gamma_1)'$, $u_1 = (u_{11}, \cdots, u_{1g})'$, y_2, θ_2, u_2 are similarly defined, and $Z = (X : l)$ with $X = (x_1', \cdots, x_g')'$ and $l = (1, \cdots, 1)'$.

Let Ω_1 be the variance of u_1 evaluated at $\hat{\alpha}_{1j}$. Similar definitions also apply to Ω_2, therefore

$$\Omega_1 = \text{diag}\left[\frac{1-\hat{\alpha}_{1j}}{n_j\,\hat{\alpha}_{1j}(\log\ \hat{\alpha}_{1j})^2}\right]_{j=1}^{g}$$

and

$$\Omega_2 = \text{diag}\left[\frac{1-\hat{\alpha}_{2j}}{n_{1j}\,\hat{\alpha}_{2j}(\log\ \hat{\alpha}_{2j})^2}\right]_{j=1}^{g},$$

where diag $[q_j]_{j=1}^{g}$ means a $g \times g$ diagonal matrix having q_j as the jth diagonal element.

By noting that the same β appears in θ_1 and θ_2, we can further combine these two equations (8) and (9) into a single equation system

$$\bar{y} = \bar{Z}\bar{\theta} + \bar{u}, \tag{10}$$

where $\bar{y} = (y_1' : y_2')'$; and $\bar{Z}$ is

$$\overline{Z} = \begin{pmatrix} X: & l: & 0 \\ X: & 0: & l \end{pmatrix},$$

with 0 a $g \times 1$ vector of zeros, $\overline{\theta}=(\beta':\gamma_1:\gamma_2)'$, and $\overline{u}=(u_1':u_2')'$. Note that the error term $\overline{u}$ is heteroscedastic. The variance matrix of $\overline{u}$ can be consistently estimated by

$$\overline{\Omega}=\text{diag}[\Omega_1:\Omega_2].$$

The minimum chi-square estimator is obtained from the weighted least squares applied to the above equation system

$$\hat{\overline{\theta}} = (\overline{Z}'\,\overline{\Omega}^{-1}\overline{Z})^{-1}\overline{Z}'\,\overline{\Omega}^{-1}\overline{y} = \overline{\theta} + (\overline{Z}'\,\overline{\Omega}^{-1}\overline{Z})^{-1}\overline{Z}'\,\overline{\Omega}^{-1}\overline{u}. \tag{11}$$

And its variance can be consistently estimated by

$$\text{var}(\hat{\overline{\theta}}) = (\overline{Z}'\,\overline{\Omega}^{-1}\overline{Z})^{-1} = \begin{pmatrix} X'\Omega_1^{-1}X+X'\Omega_2^{-1}X & X'\Omega_1^{-1}l & X'\Omega_2^{-1}l \\ l'\Omega_1^{-1}X & l'\Omega_1^{-1}l & 0 \\ l'\Omega_2^{-1}X & 0 & l'\Omega_2^{-1}l \end{pmatrix}^{-1}. \tag{12}$$

By using an inverse formula for partitioned matrices (see, for example, Amemiya 1985, p. 460), we can separate the regression coefficient estimator $\hat{\beta}$ out of $\hat{\overline{\theta}}$ in (11):

$$\hat{\beta} = [X'(\Omega_1^{*-1}+\Omega_2^{*-1})X]^{-1}[X'\Omega_1^{*-1}y_1+X'\Omega_2^{*-1}y_2], \tag{13}$$

where Ω_j^{*-1} $(j=1, 2)$ is defined as

$$\Omega_j^{*-1} = \Omega_j^{-1} - \Omega_j^{-1}l(l'\Omega_j^{-1}l)^{-1}l'\Omega_j^{-1}. \tag{14}$$

Accordingly, its variance is

$$\text{var}(\hat{\beta}) = [X'(\Omega_1^{*-1}+\Omega_2^{*-1})X]^{-1}, \tag{15}$$

which is the $k \times k$ north-west block of the ($\hat{\overline{\theta}}$) matrix in (12).

B. *Specification Test*

Besides the computational advantage of the suggested estimators, they offer a convenient way of testing the proportionality assumption in PHM. We can easily test whether the regression coefficient β in (8), say $\beta^{(1)}$, is equal to β in (9), say $\beta^{(2)}$. By running separate weighted least squares, we obtain the point estimates, $\hat{\beta}^{(1)}$ and $\hat{\beta}^{(2)}$, and their variance estimates, $\hat{\text{var}}(\hat{\beta}^{(1)})$ and $\hat{\text{var}}(\hat{\beta}^{(2)})$. These two estimators are asymptotically uncorrelated, yielding zero covariance. Therefore, $\hat{\text{var}}(\hat{\beta}^{(1)} - \hat{\beta}^{(2)}) = \hat{\text{var}}(\hat{\beta}^{(1)}) + \hat{\text{var}}(\hat{\beta}^{(2)})$.

Under the null hypothesis

$$Z=(\hat{\beta}^{(1)}-\hat{\beta}^{(2)})'[\hat{\text{v}}\text{ar}(\hat{\beta}^{(1)})+\hat{\text{v}}\text{ar}(\hat{\beta}^{(2)})]^{-1}(\hat{\beta}^{(1)}-\hat{\beta}^{(2)}), \tag{16}$$

has an asymptotic χ^2 distribution with degrees of freedom equal to the number $(=k)$ of covariates. If this statistic turns out to be bigger than the critical point of the $\chi^2(k)$ distribution, reject the proportionality, otherwise do not reject.

By using the similar methods as before, we derive:

$$\hat{\beta}^{(j)}=(X'\Omega_j^{*-1}X)^{-1}X'\Omega_j^{*-1}y_j,\ \ j=1,\ 2, \tag{17}$$

and

$$\text{var}(\hat{\beta}^{(j)})=(X'\Omega_j^{*-1}X)^{-1},\ \ j=1,\ 2. \tag{18}$$

In addition to the overall χ^2 test in (16), we can perform individual t-tests. Under the null hypothesis of proportionality, t_i

$$t_i=\frac{\hat{\beta}_i^{(1)}-\hat{\beta}_i^{(2)}}{\hat{\sigma}_{(1-2)i}},\ \ i=1,\ \cdots,\ k, \tag{19}$$

will have an asymptotic standard normal distribution, where $\hat{\beta}_i^{(j)}$ ($i=1, \cdots, k$) is the ith element of $\hat{\beta}^{(j)}$ ($j=1, 2$), and $\hat{\sigma}_{(1-2)i}$ is the square root of the ith diagonal element of $\hat{\text{v}}\text{ar}(\hat{\beta}^{(1)})+\hat{\text{v}}\text{ar}(\hat{\beta}^{(2)})$. The advantage of individual t-tests is to identify those covariates which exhibit non-proportional effects on the hazard rate, and to offer the direction of those non-proportional effects, that is, whether stronger at an earlier duration or later.

IV. Application to the Analysis of CEO Turnovers

We apply the minimum chi-square method to the analysis of the duration of CEO tenure in large U.S. corporations. By comparing the estimated coefficients of two different tenure periods, we test the proportional hazard assumption in CEO turnovers.

A. Sample and Data

The firms in our sample are mainly the NYSE-listed firms among 1991 Standard & Poor's 500. A secondary source is the Forbes magazine's annual list of CEOs. NYSE-listed firms which appeared in this list through 1986 to 1990 were added to the sample. The information about the CEOs is drawn from Dun and Bradstreet's reference Book of Corporate Managements, the Forbes' list, and

Standard & Poor's Register of Corporations, Directors and Executives. As a measure of the firm performance, we use the stock return data taken from CRSP tapes. After excluding the firms with incomplete data, the final number of firms in the sample is 454. For these firms, we observed the CEOs whose tenure ended between 1981 and 1990, inclusively.

Our interest is in identifying the termination of CEO tenures for disciplinary reasons and in studying its relationship with firm performance. For this reason, if a CEO voluntarily resigns or takes a normal retirement, we do not regard his tenure as terminated. Instead, we view it as right-censored. To identify the reasons of CEO turnovers, we consulted the Wall Street Journal Index for each turnover. When no explicit reason was available, however, the CEO's age was used as the primary criterion for determining whether the turnover was disciplinary or not. The turnovers which occurred at the age of sixty-four or more were regarded as normal retirements, and others were regarded as involuntary (disciplinary). In many companies, sixty-five is the age of normal retirement, and many top managers do retire around that age.[1]

Table 1 classifies the CEO turnovers by the reasons. Normal retirement is the most frequent reason of CEO turnovers. Very few turnovers are explicitly attributed to internal pressure from the board or large shareholders. Many turnovers are without specific reasons and are to be judged only by the age of the departing CEO. This indicates the limitation of our classification of disciplinary and non-disciplinary turnovers. Many turnovers classified as disciplinary due to lack of information may have been in fact voluntary, and others described as normal retirements may not have been as amicable as they appeared. We can also imagine a turnover which is partly disciplinary and partly voluntary. A CEO may voluntarily leave the position amid differences with the board, although he could retain his position if he insisted. To control for this voluntary aspect, we include the CEO's age (measured in years) in the estimation as a control variable in addition to firm performance measures. We expect that a manager has less incentive to hold on to the position as he approaches the retirement age.

[1]According to Yungsan Kim (1996), the percentage of departing CEOs sharply increases at the age of sixty-four and sixty-five, and remains relatively high thereafter.

TABLE 1
REASONS OF CEO TURNOVER

Reason	Frequency
Normal Retirement (or so-presumed)	168
Death or Health Problem	11
Take a Better Appointment Elsewhere	1
Resign under Pressure from Inside	11
Resign under Pressure from Outside	4
Resign amid Criticism	4
Persue Other Interest	4
Resign before Age 64 without Specific Reason	123
Pre-determined Transition	2
Others	7
TOTAL	335

B. Estimation Results

Each CEO's tenure is grouped into four intervals; less than three years, three to five years, six to nine years, and more than nine years. Out of these four intervals, we estimate the effect of performance and age on disciplinary CEO turnovers in three-to-five (hereafter, interval I) and six-to-nine (interval II) years of tenure. The first two years are not considered because a CEO is seldom dismissed due to poor performance within two years of appointment. This must be due to the fact that the firm performance in the early years of a new CEO's tenure reflects more of his predecessor's performance than his own. Also the board would allow a new CEO a couple of years of grace period to adjust to the new position.[2] We also do not consider the last interval, over nine years of CEO tenure. It is because only a small proportion of CEOs remains in the position past nine years, and when they do, they mostly resign in normal retirements.

To apply the minimum chi-square method introduced in the previous section, we categorize the age and the performance covariates used in the model. Then, for each interval, we classify CEOs into nine (three by three) different groups based on three age values (YOUNG, middle, and OLD) and three performance values

[2]See Yungsan Kim (1996).

(GOOD, medium, and BAD). Age is measured at the beginning of the interval, and firm performance is measured by average annual stock return.

YOUNG takes value one if the age is less than or equal to fifty-five, and zero otherwise. OLD takes value one if the age is over sixty. We used the middle age as the reference age group. The rationale for using the cut-off ages of fifty-five and sixty is social and biological. As mentioned before, sixty-five is the age of normal retirement in many companies. Our classification is based on whether the CEOs are more than five or ten years away from the normal retirement age. In terms of age category, the number of CEOs decreases from YOUNG to OLD.

GOOD and BAD take value one if the stock return is more than eight percent points below and above the market average, respectively, and zero otherwise. We used the medium return as the reference performance group. Noting that the standard deviation of the distribution of the excess stock returns is sixteen percent, the eight percent point criterion is chosen so that the stock returns in the one standard deviation range around the market return should form the reference performance group. As can be seen in Table 2, the CEOs in the medium performance group comprise about 40 percent. These dummy variables comprise the covariates in our estimation. For the dependent variable, we calculate the empirical turnover probability for each group and transform it as in equations (6) and (7).[3]

Table 2 shows the number of turnovers relative to the number of CEOs for each of the nine groups and for each of the two intervals. The first fraction in each cell corresponds to interval I, and the second fraction to interval II. Note that the sum of the numbers in the denominator (915) is greater than the number of the firms in the sample (454) because a CEO can appear more than once in the denominator. For example, if a CEO first assumed the office in 1979 and resigned in 1988, he contributes one to the denominator in both intervals I and II. Also, there might be more than one CEOs from a firm if the firm experienced a CEO turnover during

[3]There is one group with zero turnover. Since the dependent variable is not defined with zero cell probability, we substituted one, the smallest natural number, for zero. This is rather arbitrary but does not change the order of the probabilities among the groups, and it is not expected to seriously undermine the empirical results.

TABLE 2

RELATIVE FREQUENCIES OF CEO TURNOVER BY AGE, PERFORMANCE AND INTERVAL

Age\Performance	BAD		MEDIUM		GOOD	
YOUNG	8/61	1/26	6/96	3/60	0/91	2/37
MIDDLE	7/48	8/24	5/60	9/62	2/58	7/37
OLD	4/33	8/31	9/57	5/42	3/48	4/44

Note: Ratios represent the empirical hazard rate. The denominator is the number of all observations, and the numerator is the number of observations experiencing turnovers. In each cell, the first ratio applies to three-to-five years of tenure, and the second ratio to six-to-nine years of tenure.

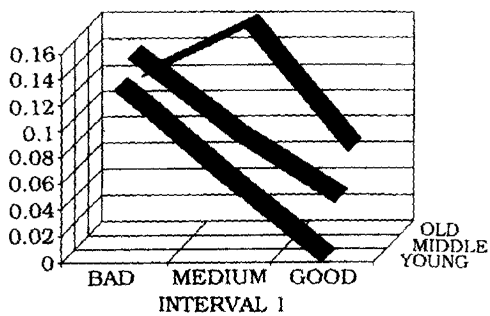

FIGURE 1

RELATIVE FREQUENCIES OF CEO TURNOVER

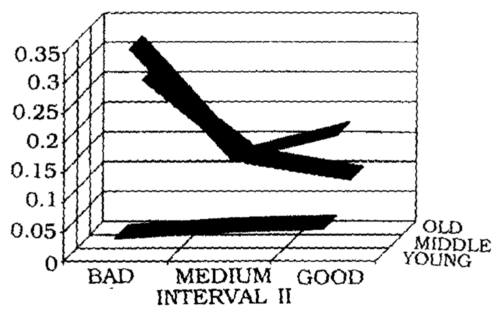

FIGURE 2

RELATIVE FREQUENCIES OF CEO TURNOVER

the sample period. On the other hand, when a CEO's tenure is right censored, that is, when the CEO leaves the firm for a non-disciplinary reason at some point in an interval, the CEO is excluded from calculating the number in the denominator of the corresponding interval. Also, since we do not consider the CEO tenures outside the two intervals, some of the CEO turnovers in Table 1 are not included in Table 2.

Figure 1 and Figure 2 show the relative frequencies of disciplinary CEO turnovers in intervals I and II, respectively. By comparing these two figures, one can see different age and performance effects on CEO turnovers across the two different intervals. These disparate covariate effects will be formally tested shortly.

Table 3 shows the estimated coefficients for the two intervals and their differences. In interval I, the age dummies are not statistically significant though the signs indicate an increasing probability of turnovers with age. The good performance dummy has a negative and significant coefficient. In interval II, YOUNG has a negative and significant coefficient, and BAD has a positive and significant coefficient. The difference of the coefficients of YOUNG and GOOD between the two intervals is significant at the 10 percent level. The chi-square statistic for the hypothesis that the two sets of coefficients are the same across the two intervals is 8.75, which is significant at 10 percent level.

The negative coefficient of YOUNG in the interval of six to nine years of tenure seems to reflect the voluntary aspect of the turnovers which we might have mistakenly regarded as disciplinary. This voluntary aspect is expected to increase with a CEO's age and tenure. The relatively higher probability of turnovers for middle-to-old CEOs in interval II should be due to the effect of higher age and the resulting higher proportion of normal retirement.

The differences in the effects of performance across the two intervals reveal an interesting nature of CEO turnovers. In both intervals, the risk of a turnover significantly increases as the performance changes from GOOD to BAD. Though the increase is greater in interval I than in II, the difference is not significant. What is significant is the difference in the increase of the risk from good to medium performance. For the CEOs of three to five years of tenure, average CEOs are not much better than those with very bad performance in terms of the turnover risk. Performance matters only when it is very good. However, when the tenure is six to nine

TABLE 3
MINIMUM CHI-SQUARE ESTIMATION OF CEO TURNOVER

Covariates	All	3~5years I	6~9years II	Difference I-II
Base line hazard I	−2.17**	−2.33**		
	(−8.78)	(−7.02)		
				−0.58
				(−1.34)
Base line hazard II	−1.74**		−1.75**	
	(−7.55)		(−6.25)	
YOUNG	−0.75**	−0.26	−1.40**	1.14*
	(−2.66)	(−0.71)	(−3.05)	(1.92)
OLD	−0.05	0.41	−0.39	0.80
	(−0.23)	(1.10)	(−1.21)	(1.62)
BAD	0.59**	0.42	0.83**	−0.41
	(2.51)	(1.27)	(2.37)	(−0.86)
GOOD	−0.38	−1.07**	0.08	−0.99*
	(−1.34)	(−2.28)	(0.21)	(−1.92)
Overall χ^2-test (d.f.=4)				8.75*

Note: The numbers in the second column are the estimated coefficients under the proportionality assumption using the data from interval I and II. The third and fourth columns include the estimated coefficients of each interval. Those in the fifth column are their differences. YOUNG (OLD) is a dummy variable indicating whether the CEO's age is no more than fifty-five (more than sixty). GOOD (BAD) is a dummy variable indicating whether the average stock return of the firm is eight percent above (below) the market average. Therefore, the reference age group is fifty-six to sixty, and the reference performance group is negative eight percent to positive eight percent excess return. The numbers in parentheses are asymptotic t-ratios.

**: significant at 5% level

* : significant at 10% level

years, an average CEO's turnover risk is not different from those with very good performance. At a later tenure period, it is only when the performance is very bad that the turnover risk increases significantly. To sum, a CEO with short tenure faces a higher risk of disciplinary turnover until he proves his ability, whereas one with longer tenure enjoys a lower risk unless the performance is really bad, creating asymmetric disciplinary effects of the CEO performance across the tenure periods.

We propose two reasons for the above asymmetric effects. First, the differences in average CEO quality across the two periods cause the asymmetric effects. Long-tenured CEOs may have, on average, better reputation than their short-tenured counterparts. This is possibly due to unobserved heterogeneities among CEOs. Inferior CEOs are more likely to be replaced earlier in their tenure. Thus, surviving CEOs are more likely to have good unobserved attributes, unobserved only to the econometricians. As the directors can recognize these good attributes of the survivors, the CEOs in later tenure period are not likely to be dismissed unless their recent performance is extremely poor.

Second, it is plausible that the CEOs with longer tenure are more entrenched. It may be either because they have gained more political power in the firm or because the firm has become more dependent upon the CEO.[4] Disentangling these two effects invokes an identification problem, and is left for future research.

V. Concluding Remarks

Often, we face grouped duration data due to discrete nature inherent in much of survey design. The proportional hazard model is the most widely used continuous time duration model. When covariates are all discrete or categorical, application of Berkson's minimum chi-square estimation yields computationally very simple estimators which are asymptotically as efficient as the maximum likelihood estimators. Our estimation method allows a straightforward generalization when there are more than two intervals in the discrete observation scheme.

The suggested estimation method and the specification test are applied to the analysis of CEO turnovers in larges U.S. corporations over the period 1981 to 1990. We found that elderly CEOs are more likely to be turned over and that better-than-average performers tend to stay longer. We also observed an asymmetry in the effect of performance on disciplinary CEO turnovers. At earlier tenure as CEO, good performance matters, increasing the chance of survival.

[4]For a theoretical discussion of managerial entrenchment, see Schleifer and Vishny (1989).

On the other hand, at later tenure, bad performance makes a difference, increasing the possibility of turnovers.

Regarding the asymmetric performance effect on the disciplinary CEO turnovers, we suggested two candidate reasons. One was heterogeneity argument, saying that CEOs at an earlier stage are on the average less qualified than the CEOs who have already survived the initial test of the market in terms of unobserved talents. The other was entrenchment argument, claiming that the CEOs in their later tenure periods are better protected from the recent poor performance than the CEOs at an earlier stage.

By introducing unobserved heterogeneity into the model, we expect to sort out the above two arguments to a certain extent. It would be interesting to know how much of the asymmetry is caused by the unobserved heterogeneity and how much by the entrenchment, which is left for future research.

(Received August, 1998; Revised September, 1998)

References

Amemiya, T. *Advanced Econometrics*. Cambridge: Harvard University Press, 1985.

Berkson, J. "Application of the Logistic Function to Bio-Assay." *Journal of the American Statistical Association* 39 (1944): 357-65.

Cox, D. R. "Regression Models and Life-Tables." *Journal of the Royal Statistical Society B* 34 (1972): 187-220.

__________. "Partial Likelihood." *Biometrica* 62 (1975): 269-79.

__________, and Oakes, D. *Analysis of Survival Data*. London: Chapman & Hall, 1984.

Furtado, E. T., and Karen, V. "Causes, Consequences and Shareholder Wealth Effects of Management Turnover: A Review of the Empirical Evidence." *Financial Management* 19 (1990): 60-75.

Harrison, J. R., Torres, D., and Kukalis, S. "The Changing of the Guard: Turnover and Structural Change in the Top-Management Positions." *Administrative Science Quarterly* 33 (1988): 211-32.

Jensen, M. C., and Murphy, K. "Performance Pay and Top-Management Incentives." *Journal of Political Economy* 98 (1990): 225-64.

Kalbfleisch, J. D., Prentice, R. L. *The Statistical Analysis of Failure*

Time Data. New York: John Wiley & Sons, 1980.

Kiefer, N. M. "Analysis of Grouped Duration Data." *Contemporary Mathematics* 80 (1988): 107-37.

Kim, Yungsan. "Long-Term Firm Performance and Chief Executive Turnover: An Empirical Study of the Dynamics." *Journal of Law Economics and Organization* 12 (1996): 480-96.

Lee, Lung-fei. "Semi-parametric Minimum Distance Estimation." CREST Working Paper, No. 92-08, Department of Economics, University of Michigan, 1992.

Prentice, R., and Gloeckler, L. "Regression Analysis of Grouped Survival Data with Application to Breast Cancer Data." *Biometrics* 34 (1978): 57-67.

Ryu, K. "Group Duration Analysis of the Proportional Hazard Model: Minimum Chi-square Estimators and Specification Tests." *Journal of the American Statistical Association* 89 (1994): 1386-97.

Schleifer, Andrei, and Vishny, R. W. "Managerial Entrenchment: The Case of Manager-Specific Investments." *Journal of Financial Economics* 25 (1989): 123-39.

Sueyoshi, G. T. "Evaluating Simple Alternatives to the Proportional Hazard Model: Unemployment Insurance Receipt and the Duration of Unemployment." Mimeograph, University of California, at San Diego, 1991.

Warner, J. B., Watts, R. L., and Wruck, K. H. "Stock Prices and Top Management Changes." *Journal of Financial Economics* 20 (1988): 461-92.

Zellner, A. "An Efficient Method of Estimating Seemingly Unrelated Regression Equations and Tests of Aggregation Bias." *Journal of the American Statistical Association* 57 (1962): 500-9.

________, and Lee, T. "Joint Estimation of Relationships Involving Discrete Random Variables." *Econometrica* 33 (1965): 383-94.