UNIVERSITY OF NAIROBI

COLLEGE OF BIOLOGICAL AND PHYSICAL SCIENCES

SCHOOL OF MATHEMATICS

PROJECT IN SOCIAL STATISTICS

MODELLING
TEACHERS' DURATION IN SERVICE
IN PUBLIC SCHOOLS AND INSTITUTIONS IN KENYA

A PROJECT SUBMITTED TO THE SCHOOL OF MATHEMATICS IN PARTIAL FULFILLMENT

FOR A DEGREE OF MASTER OF SCIENCE IN SOCIAL STATISTICS

©JULY 2008

Declaration

I, the under signed, declare that this project is my original work
and to the best of my knowledge has not been presented for the
award of a degree in any other University.

MWAMBI JERITA JEMIMAH
Reg. No I56/7899/2006

Signature. . . . . Date. . . 12/9/2008.

Declaration By Supervisor

This project has been submitted for examination with my approval as supervisor

Dr. Rosemary Nguti

School of Mathematics,

University of Nairobi,

P.O. Box 30197, Nairobi.

Signature. . . . . Date. . 22/9/2008. . . . .

i

Acknowledgments.

Dedicated to:

To my husband, Morgan Kiseu for your continued support and encouragement;

My children Bryan, Steve and sweet Brenda for your understanding.

And to my mum for your moral support and believing in me.

Abstract

There has been increased enrollment in both primary and secondary schools in Kenya since the government implemented the free primary education in 2003 and now the subsidized secondary education. This means a higher demand for qualified teachers, hence the need to retain the already hired, trained and experienced teachers. In order to achieve this, there's need to investigate the factors associated with teachers duration in service. This study used survival analysis methods and in particular the Cox proportional hazard model to determine a combination of potential explanatory variables that are associated with teachers' duration in service for teachers employed by the Teachers Service Commission.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Education is a very powerful tool for economic growth in a developing country like Kenya. For Kenya to develop industrially and technologically by 2020, education will have to be improved. There has been increased enrollment in both primary and secondary schools. In public primary schools, the enrollment has risen by 1.5 million children in less than one year, since the government launched and implemented the free primary education (FPE) programme in January 2003.

According to the Economic survey 2007, total enrollment of pupils in standard 1 to 8 in public primary schools was 7.63 million in 2006, up from 7.59 million in 2005; while in secondary schools enrollment increased from 934,149 in 2005 to 1,030,080 in 2006. This will mean a demand in recruitment of more teachers and retaining them in profession. One of the major challenges in improving the quality of education in this nation is the rising attrition rate of teachers. The total number of teachers in service in 2003 was 227,467, while in 2006 was 207,505;the number of primary teachers dropped from 171,033 in 2005 to 162,993 despite the government recruiting 5,641 primary school teachers in 2006. The number of teachers in secondary schools went down by 10.6% from 47,435 in 2005 to 42,403 in 2006. It is estimated that 1.8 % of the country's teachers are

1

dying annually and still others leave through resignation, transfer of service and dismissal.

In Kenya the government is the main employer of teachers through the Teachers service commission (TSC) which was established in 1967 by an act of parliament (cap 212 of the laws of Kenya), to provide services to teachers. The TSC is mandated the management function of registration, recruitment, deployment, renumeration, promotion, and discipline of teachers. Due to the need to contain the wage bill strategy, the TSC has only been replacing teachers who exit service through natural attrition.

Educational researchers have long been interested in investigating teacher attrition and factors associated with it. Teacher attrition according to Boe, Bobbit, and Cook (1993), is a component of teacher turnover (i.e, changes in teacher status from year to year). Teacher turnover may include teachers leaving the profession, but may also include teachers who change fields or schools. The rates of attrition depend on this definition.

1.1 Problem statement

Investigation of teachers' duration in service and the factors associated with it, is of great importance. Whether attrition of teachers is due to natural causes or not, causes the government to incur huge losses in replacement of these teachers by other newly trained ones. It is costly and the wage bill for teachers is already over stretched. There's need to retain the already hired, trained and experienced teachers as the government cannot afford to lose such teachers through voluntary termination. Teachers' termination from service therefore needs to be addressed if the country is to attain education for all(EFA) by 2015 and enhance equitable and quality education in primary and secondary schools. This study will explore the factors associated with termination of teachers' service. There's need therefore to investigate termination of teacher's service and the factors associated with it. There's need to go beyond the simple annual attrition rates in research on teacher

with it. There's need to go beyond the simple annual attrition rates in research on teacher career paths; this study will use survival analysis procedures which are able to reveal much more about teacher's exit from service.

1.2 The general and specific objectives of the study

The overall aim is to determine factors associated with teachers' duration in service. The specific objectives of this study are:

1. To find out how long an average teacher can stay in the classroom before exit.

2. To determine the differences in time to termination by types of exit.

3. To determine correlates of termination from service.

1.3 Research questions

1. How long can an average teacher stay in the classroom before exit?

2. Are there any differences in time to termination by types of exit?

3. What are the correlates of termination from service?

1.4 Significance of the study

Teachers' exit from service is still a major challenge in this country. Teachers' termination from service therefore needs to be addressed if the country is to attain education for all(EFA) by 2015 and enhance equitable and quality education in primary and secondary

schools. This study will explore the factors associated with termination of teachers' service.

This study will provide more insight into this issue of termination of teachers' service since it will use methods of survival analysis to explore the effects of factors associated with the exit of teachers from service with the survival experience of teachers.

The findings from this study will provide a lot of information to the educational administrators, planners, and policy makers for proper planning.

This study will also contribute to the body of knowledge on value addition on the issue of factors associated with termination of teachers from service.

1.5 Motivation of the study

The motivating factor of this study is the rising attrition rate of teachers in publc schools and institutions. The data used in this study is from the teachers service commission (TSC) given with permission from the ministry of education and the teachers service commission. The data is based on 164,282 teachers who have been in service to early 2008, including those who have exited either through death,retirement, transfer of service,resignation and dismissal. It includes all teachers employed by the government through the teachers service commission to teach in all public primary schools and post primary institutions; these are teachers trained in public primary school teacher training colleges, diploma teacher training colleges, technical teacher training colleges and public universities and untrained teachers. Their date of hire and exit(for those who have exited) is noted; other important information for each teacher is included like age, job group, job designation, marital status, gender, working province,and nature of exit.

In this study teachers who are still in service to April 2008 and including those who had left the profession were considered. However the data availed information from 2003

for those who left the profession. The teachers were followed from their point of entry into the teaching career; though not all were recruited at exactly the same time, to the time when they exited or left the teaching profession.The teacher could have left the teaching profession through death, retirement, transfer of service, resignation, or dismissal.

The outcome of interest which is the time to event,in this study was the time to exit from the teaching profession. For each teacher his/ her survival time was given in years(These are the number of years they have worked for in the TSC) along with the censoring indicator. Those teachers who are still working were said to be censored as they had not experienced the event of interest which is leaving the profession at the time of data collection, indicated by (0); while those who left the teaching profession are not censored indicated by (1). The additional information collected for every teacher, that is: age, job group, gender, marital status, type of exit and working province form the explanatory variables or factors associated with teachers' duration in service.

Chapter 2

Literature Review

There are a number of educational researchers who have long been interested in carrying out research on teacher career paths and mainly on teacher attrition and factors associated with it. Since 1970's research shows teacher attrition to be a problem. Most of the quantitative research on teacher career paths has been descriptive focusing on determining the number of the nation's current teachers that are likely to remain in the classroom. Attrition rate is a simple estimate of the proportion of teachers employed in a specific jurisdiction (e.g a district, a province and so on), who leave their jobs during a finite period , usually one year.

Grissmer and Kirby (1987) presented annual attrition rates in Illinois, Michigan, new York and Utah culled from state reports from the period 1979 to 1982. They reported that about 6% of the teachers employed in any given year were gone the following year.

Bobbit, Faupel and Burns(1991) used the 1987-1988 schools and staffing survey and 1988-1989 teacher follow-up survey, and expressed the attrition rate for teaching profession as 5.6% in the public schools and 12.7% in private schools. Boe, Cook, Bobbit and Weber (1995) estimated attrition rates for general teachers and special education teachers , using the data from 1990-1991 schools and staffing survey and 1992 teacher follow-up

survey. They reported 6.3% of the special education teachers and 5.6% of the general education left the profession. The attrition was calculated by getting the proportion of teachers employed in special education and general education , who left their jobs after one year and expressed it as a percentage.

Survival analysis methods have been used in research on teacher career paths and has been found to reveal more information than the simple attrition rates. According to Singer and Willet (winter, 1991), academic researchers often estimate the teacher retention or survival rate, instead of estimating the teacher attrition rate. The teacher retention or survival rate is the proportion of teachers in one year, who continue to teach the next.

According to Singer and Willet (2003), the application of the "Whether and When to test" may reveal the suitability of the methodology of survival analysis for addressing various research questions. And they suggested that if one's research questions include either word- Whether or When-, then one would probably need to use survival methods. Such questions are common in educational research.

Whitener (1965) and Charters (1970) were the first to apply the concept of teacher survival rates to teaching. They used a 5-year time frame which resembles the 5-year rates used by physicians to report life expectation of patients after a surgery. Whitener examined the survival rates of 937 public school teachers in Missouri from 1951 to 1953; while charters studied 2,064 teachers who began working in Oregon in 1962. Their general results were very similar in that, a very high proportion of those who enter teaching in any given year do not return for a second year. A smaller loss takes place from the 2^d to $3^{rd} year$, and the process of attrition slows down and eventually becomes stable from year to year. Charters found that males tend to survive longer than females.

Heynes (1988) in her follow up of the National Longitudinal Study (NLS) found that 60% - 70% of the NLS sample who became teachers were still teaching (surviving) after 5 years in the classroom. She presented the 5-year attrition rates by school type and location; and concluded that teachers are more likely to leave good schools in good location than

7

problem schools in difficult locations.

Mark and Anderson (1978) examined the survival rates of public school teachers from 1968 to 1976. They looked at the survival behavior of each cohort of new entrants to the teaching profession between 1968 to 1975. They were able to examine how survival rates change over a period of time.

In a similar study to update the 1978 study, Mark and Anderson (1985) examined the lengths of employment of 14,827 teachers who were hired by districts in the st. Louis, Missouri area between 1969-1982. They tracked cohorts of new teachers for as long as 13 years, and by comparing their survival rates documented the large proportions of newly hired teachers who left after only a few years in the classroom.

Singer and Willet (winter, 1991) emphasized the use of longitudinal data as it permits a more refined and realistic view of a teacher's career; an ability to track factors associated with teacher's stay or leave decisions. Such data lends itself naturally to survival methods. They noted that, longitudinal data enables tracking of one or more cohorts of new teachers over time, noting each person's employment status along the way. Some of the educational researchers who have used longitudinal data include Charters (1970), Murnane, Singer and Willet(1988,1989),Schlechty and Vance (1981), and Grissmer and Kirby (1987).

Singer and Willet (1988) used the data presented by Mark and Anderson (1985) and used a set of techniques developed by actuaries and bio statisticians to address problems of bias brought in by ignoring right- censoring or the discrete nature of the observations. They applied the survival function, median total lifetimes, and the hazard function to the study of teacher's career paths; and were therefore able to uncover career disruptions caused by involuntary layoffs. Singer (1992) again illustrated the application of the survival function by using data describing the career paths of 3,941 special educators hired in Michigan in the 1970,s. He found that the greatest risk of leaving the profession is highest during he first few years on the job and drops in subsequent years.

Schlechty and Vance (1981) examined tables of estimated survivor functions for teach-

ers with differing test scores to show that teachers with highest test scores are more likely to leave. Murnane et al (1991) used hazard model to show that women who were younger than 30 years when hired, were more likely to leave their teaching jobs than were either older women or men of any age. By including time-varying salary as a predictor, Murnane also showed that salary is a main reason for teachers' exit from the profession; he noted that better paid teachers stay in the profession longer.

Theobald (1990) also noted that, salaries are positively related to decisions to continue teaching in the same district. In the same study he noted that decisions to continue teaching in the same district positively relate to experience.

Kirby and Grissmer (1993) used the human capital model to give reasons for teacher attrition. They noted that the individual weighs costs and benefits; such that most teachers leave teaching during the early part of the career because the teacher accumulates less specific capital(knowledge specific to occupation and that which is non-transferable). With increasing experience less teachers tend to leave the profession because more specific capital exists.

Marital status is another factor associated with the teachers' decision to stay or leave the profession. According to Grissmer and Kirby (1991) theorized that the decision to accept and keep a teaching job depends on life cycle factors (existing family status and change in family status). Bloland and Selby (1980) note that the earlier research indicates that preference of the spouse leaving or staying in the teaching profession is one of the most important factors for staying in the field of education.

Chapter 3

Methodology

3.1   Introduction of survival analysis

Survival analysis is a collection of statistical procedures for data analysis for which the outcome variable of interest is time until an event occurs; or in other words survival analysis is concerned with analysis of times from some specific time origin until some critical event or end-point of interest. In employment context, this may be time to exit from employment, and this could be through death, resignation, dismissal or retirement. Another example of time to event modeling could be the rate or time to which former convicts commit a crime again after they've been released. In this case, the 'event' of interest would be committing a crime after release. The start or end points for measuring times are chosen according to the context and may not represent the entire life of an individual.

Survival analysis attempts to answer questions such as: what is the fraction of a population which will survive past a certain time? Of those that survive at what rate will they die or fail? Can multiple causes of death or exit in this case be taken into account? How do particular circumstances or characteristics increase or decrease the risk of event?.

Survival analysis provides a straight forward easily application tools for appropriately analyzing the timing of educational transitions i.e. time from entry to exit from career.

The survival time of an individual is censored when the end-point of interest has not been observed for that individual or the survival status of an individual at the time of the analysis might not be known because that individual has been lost to follow-up.As an example, a patient may move to another part of the country , or to a different country after being recruited to a clinical trial and can no longer be traced. The only information available on the survival experience of that patient is the last date on which he or she known to be alive. This may be the last date the patient reported to a clinic for regular check-up. When death of a patient is from a cause unrelated to the treatment , then the actual survival time can also be regarded as censored. A good example is a patient in a clinical trial to compare alternative therapies for prostatic cancer who dies in a tragic road accident. However if the accident resulted from an attack of dizziness, which might be a side- effect of the treatment to which that patient had been assigned, then the death is not unrelated to the treatment. In this study the event may probably not have occurred for some teachers at the end of the follow-up period, that is, those teachers who do not exit teaching will possess censored event times. It is not known when or whether the teacher will experience the event; all we know is that by the end of the data collection, the event has not yet occurred.

An individual's exact survival time may become incomplete at the right hand side of the follow-up period which may occur when the study end or when the person is lost to follow-up. This kind of censoring is known as right censoring. Other types of censoring are left censoring and interval censoring, right censoring is the most common in survival data. Left censoring is encountered when the actual survival time of an individual is less than observed. A good illustration of this form of censoring is to consider a study in which the interest is mainly on the time to recurrence of a particular cancer following surgical removal of the primary tumor. The patients are examined to determine if the cancer has

11

recurred three months after the operation. Some of the patients may be found to have a recurrence, the actual time to recurrence is less than three months and the recurrence times of these patients is said to be left censored. In interval censoring, individuals are known to have experienced a failure within an interval of time. If in the above example the patient is observed to be free of the disease at three months, but is found to have had a recurrence when examined six months after surgery, the observed recurrence time is then said to be interval-censored.

3.2 Survival Data

In the presence of censoring survival data consists of pairs of observations (t_i, δ_i) where i denotes the subject ranging from 1 to n. The presence or absence of censoring is denoted by the δ_i where,

$$\delta_i = \begin{cases} 1, & for\,Ending\,Event\,Observed \\ 0, & for\,observation\,Censored \end{cases}$$

The survival time is denoted by t_i, the time till the event and the time until censoring occurs if the observation is censored. Thus, one sample of survival data contains n pairs of survival times and censoring indicators.

3.3 Survival function and Hazard function

The most commonly used tools for analyzing time to event data are the survival function and hazard function. The methods for estimating survival function and hazard function are said to be non-parametric or distribution-free, since they do not require specific assumptions to be made about the underlying distribution of the survival times.

3.3.1 The survival function

The survival/survivor function is the probability of surviving beyond sometime t (i.e. experiencing the event after sometime t). It is denoted by S(t), which is defined as:

$$S(t) = Pr\{T > t\}$$

where T is a random variable denoting the time of event and t an arbitrary time point. The most commonly used estimator of the survivor function is the Kaplan Meier estimator and can be used when there are censored observations.

In general, suppose there are n individuals with observed survival times $t_1, t_2, t_3, \ldots, t_n$. Some of these observations may be right-censored, and there may be more than one individual with the same observed survival time. We suppose that there are r death times amongst the individuals, where $r \leq n$. After arranging these death times in ascending order, the jth is denoted by t_j, for $j = 1, 2, \ldots, r$, and so the r ordered death times are $t_{(1)} < t_{(2)} < \ldots t_{(r)}$. The number of individuals who are alive just before time $t_{(j)}$, including those who are about to experience the event at this time, will be denoted by n_j known as the risk set at $t_{(j)}$, for $j = 1, 2, \ldots, r$, and d_j will denote the number who experience the event at this time. The time interval from $(t_{(j)} - \delta, t_{(j)})$, where δ is an infinitesimal time interval, then includes one death time. since there are n_j individuals who are alive just before $t_{(j)}$ and d_j deaths at $t_{(j)}$, the probability that an individual experience the event during the interval from $(t_{(j)} - \delta, t_{(j)})$ is estimated by $\frac{d_j}{n_j}$. The corresponding estimated probability of surviving through that interval is then $\frac{(n_j - d_j)}{n_j}$. We assume that events occur independently.

It sometimes happens that there are censored survival times which occur at the same time as one or more deaths, so that a death time and a censored survival time occur simultaneously for different individuals. In this case the censored survival time is taken to occur immediately after the death time when computing the values of the n_j. The interval from $(t_{(j)}, t_{(j+1)} - \delta)$ contains no event, therefore the probability of surviving

from $(t_{(j)}, t_{(j+1)})$ is 1. Assuming that the deaths of individuals in the sample occur independently of one another, then the estimated survivor function at any time in the k-th time interval from $(t_{(k)}, t_{(k+1)})$, k=1,2,...,r, where $t_{(r+1)}$ is defined to be ∞, will be the estimated probability of surviving beyond time $t_{(k)}$. The estimated survival function is then given by the Kaplan Meier estimator :

$$\hat{S}(t) = \prod_{t_{(j)}=t}^{k} \left[\frac{n_j - d_j}{n_j}\right] \tag{3.1}$$

where n_j is the number of individuals at risk at time $t_{(j)}$, d_j is the number of events at time $t_{(j)}$. A plot of the Kaplan-Meier estimate of the survivor function is a step-function, in which the estimated survival probabilities are constant between adjacent event times and decrease at each event time. The Kaplan-Meier estimate is also known as the product-limit estimate of the survivor function.

In this study the survival function will be used to summarize the survival experience of teachers in this study. The survival function then will be the probability that a teacher will remain employed (survive) for a least t years, T denotes employment duration.Because no more teachers can survive through year t than survived through year (t-1), S(t) is monotonically non-increasing and usually a decreasing function of time. By plotting the survival function of the teachers depending on the covariate available we will be able get a lot of information on effects of age, marital status, working province of the teacher on the survival experience of the teachers. In our study we will plot the survival function of the teachers depending on the covariate exit name.

The median time to event i.e the smallest time point such that $S(t_{MD}) \approx 0.5$, which in this case is the time beyond which 50% (half) of the teachers are expected to survive or rather be still employed. This estimation answers the question of how long an average teacher can stay in the profession before leaving. The hazard function will help us estimate the risk of a teacher leaving per unit time, e.g a year, given that he or she has survived until the beginning of that year.

14

3.3.2 Estimating the median of survival times

The median is a preferred summary measure of the location of the distribution, since the distribution of survival times tends to be positively skew. Once the survivor function has been estimated, it is easy to obtain an estimate of the median survival time. This is the time beyond which 50% of the individuals in the population under study are expected to survive, and is given by that value $t(50)$ which is such that $S\{t(50)\} < 0.5$.

Since the non-parametric estimates of S(t) are step-functions, it may not be possible to realize an estimated survival time that makes the survivor function exactly equal to 0.5. Instead the estimated median survival time, $\hat{t}(50)$, is defined to be the smallest observed survival time for which the value of the estimated survivor function is less than 0.5. Since the estimated survivor function only changes at an event time, this is equivalent to the definition

$$\hat{t}(50) = min\{t_{(j)}|\hat{S}(t_{(j)})) \leq 0.5\} \tag{3.2}$$

where $t_{(j)}$ is the j'th ordered event time, $j = 1, 2, \ldots, r$.

3.3.3 Comparison of survival curves

A common goal in survival analysis is to determine if groups of subjects have different survival characteristics.The most commonly used test is the log-rank which can be used to quantify the extent of between-group differences.

The log-rank test for differences across g groups is based on the statistic

$$U_T = \Sigma_{k=1}^{g} w_k(d_k - e_k)$$

where w_k is a code assigned to the k'th group, $k = 1, 2, \ldots, g$, d_k and e_k are the observed and expected numbers of death in the k'th group. The resulting statistic formed is

$$U_L'V_L^{-1}U_L$$

then has a chi-squared distribution on (g-1) d.f. under the null hypothesis of no differences across the g groups, V is the covariance matrix.

3.3.4 The hazard function

The hazard function is a risk function or the instantaneous risk or failure rate. The hazard function is the probability that an individual fails at time t, conditional on having survived to that time and it is defined by:

$$h(t) = lim_{\Delta t \to 0} Pr \left[\frac{t \leq T \leq t + \Delta t / T \geq t}{\Delta t} \right] \tag{3.3}$$

If T is continuous then,

$$h(t) = \frac{f(t)}{S(t)} \tag{3.4}$$

where f(t) is the probability density function and S(t) is the survival function. It then follows that

$$f(t) = -\frac{d}{dt} \{ log S(t) \} \tag{3.5}$$

and

$$S(t) = exp\{-H(t)\}, \tag{3.6}$$

where

$$H(t) = \int_0^t h(u) du. \tag{3.7}$$

is the cumulative hazard function. Thus knowledge of the basic function leads to the derivation of the others.

3.4 Modelling Survival data

In addition to observing the time to event; other characteristics may also be observed such as gender, age, heart rate, marital status and many others of the study subject. Hence

16

there's need to take into account the effect of these factors on the time to event or on the risk of the event. Through a modelling approach to the analysis of survival data, we can explore, for example, how the survival experience of a group of patients depends on the values of one or more explanatory variables, whose values have been recorded for each patient at the time origin. In the analysis of survival data, interest centers on the risk or hazard of death at any time after the time origin of the study. Consequently, the hazard function is the modelling tool in survival analysis. There are two broad reasons for modelling survival data. One is to determine a combination of potential explanatory variables that affect the risk of event. Another reason for modelling the hazard function is to obtain an estimate of the hazard function itself for an individual.

3.4.1   Cox regression model

The most commonly used hazard model in survival analysis is the Cox regression model. It allows us to determine a combination of potential explanatory variables that affect the risk of event and to obtain an estimate of the risk function. For example, in constructing a model of length of employment, variables such as age, gender, job category could be included in the model. By using cox regression also known as Cox Proportional hazard model, estimated coefficients for each of the covariates is provided and allows the assessment of the impact of multiple covariates in the same model.

If the hazard of event at a particular time depends on the values $x_1, x_2, \ldots, x_p$ of p explanatory variables $X_1, X_2, \ldots, X_p$, the values of these variables will be assumed to have been recorded at the time origin of the study. The set of values of the explanatory variables in the proportional hazards model will be represented by the vector **x**, so that $x = (x_1, x_2, \ldots, x_p)'$. Let $h_o(t)$ be the hazard function for an individual for whom the values of all the explanatory variables that make up the vector **x** are zero. The function $h_o(t)$ is called the baseline hazard function. The hazard function for the ith individual

is written as

$$h_(t) = \psi(x_i)h_o(t)$$

Where $\psi(x_i)$ is a function of the values of the vector of explanatory variables for the $i-th$ individual. The function $\psi()$ can be interpreted as the hazard at time t for an individual whose vector of explanatory variables is x_i, relative to the hazard for an individual for whom $\mathbf{x} = 0$.

Since the relative hazard $\psi(x_i)$ cannot be negative, this is conveniently written as $exp(\eta_i)$, where η_i is a linear combination of the p explanatory variables in x_i. Therefore,

$$\eta_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots, + \beta_p x_{pi} = \Sigma_{j=1}^{p} \beta_j x_{ji}$$

η_i is known as the linear component of the model, also known as the risk score for the $i-th$ individual. In general if $X_1, X_2, \ldots, X_p$ are explanatory variables of interest then, the hazard function for the i-th individual is given as:

$$h_i(t) = h_o(t)exp(\beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_p x_{pi}) = h_o(t)exp(\eta_i) \qquad (3.8)$$

where $\beta_1, \beta_2, \ldots, \beta_p$ are unknown regression parameters. This is the Cox proportional hazard (PH) model, where $h_o(t)$ is left arbitrary (unspecified).

Cox Proportional hazard model make the following assumptions;

Assumptions

- There is a baseline function $h_o(t)$ common to all individuals in all the study groups.

- Each study group has a hazard function that is a positive multiple of the baseline hazard.

- Explanatory variables act only on the hazard ratio. They do not affect the baseline hazard.

- Independence of observations.

- Sufficient data for inference.

- Censoring is independent of the event of interest.

The hazard function may depend on two types of variable, namely variates and factors. A variate is a variable which takes numerical values which are often on a continuous scale of measurement, such as age or levels of blood urea nitrogen. A factor is one which takes a limited set of values which are known as the levels of the factor, for example sex is a factor at two levels. Variates and factors, either alone or in combination, are readily incorporated in a proportional hazards model. If there two factors for example A and B, and the hazard of event depends on the combination of levels of A and B, then A and B are said to interact. The effect is known as interaction.

An important principle in statistical modelling, is that an interaction term should only be included in a model when the corresponding main effects are also present. Another type of term that might be needed in a model is a mixed term formed from a factor and a variate. Terms of this type would be used when the coefficient of a variate in a model was likely to be different for each level of a factor. Interpretation of the regression parameter depends on the type of explanatory variables. For example when interpreting the parameter of a variate like age the hazard function for the i-th individual, for whom X takes the value x_i, is

$$h_i(t) = h_o(t)exp(\beta x_i)$$

The hazard function for two individuals whose covariate value differ by one unit , say x and x+1 is:

$$h_x(t) = h_o(t)exp(\beta x)$$

and

$$h_{x+1}(t) = h_o(t)exp(\beta(x+1))$$

19

respectively. Taking the ratio of the two hazards we get,

$$\frac{h_o(t)exp(\beta x + \beta)}{h_o(t)exp(\beta x)} = exp(\beta)$$

Thus $exp(\beta)$ is a hazard ratio (HR) which is the change in risk for every unit increase in X. When interpreting parameter estimates for a factor like marital status which has 'r' levels, dummy variables are introduced. Suppose the i-th individual comes from the k-th level of the explanatory variable then

$$h_i(t) = h_o(t)exp(\alpha_k)$$

This implies that

$$\frac{h_i(t)}{h_o(t)} = exp(\alpha_k)$$

but, $h_o(t)$ is the hazard function for an individual with zero covariate-variate values implying that, it is the hazard function for an individual from the r-th level. Therefore $exp(\alpha_k)$ is the HR for an individual from the k-th level relative to that of an individual from the r-th level.

3.4.2 Fitting the proportional hazards model

Fitting the proportional hazards model given above in equation (3.8) to an observed data requires the estimation of the unknown coefficients of the explanatory variables $X_1, X_2, \ldots, X_p$ in linear component of the model, $\beta_1, \beta_2, \ldots, \beta_p$. The baseline hazard function, $h_o(t)$, may also be estimated. These two components of the model are estimated separately by first estimating the β coefficients using the method of maximum likelihood. In this method, the likelihood of the sample data is obtained first, which is the joint probability of the observed data, regarded as a function of the unknown parameters in the assumed model. For the proportional hazards model, this is a function of the observed survival times and the unknown $\beta - parameters$ in

20

the linear component of the model. Estimates of the $\beta's$ are then those values which are the most likely on the basis of the observed data. These maximum likelihood estimates are therefore the values which maximize the likelihood function.

Suppose that the data for n individuals consists of r distinct event times and n-r right-censored survival times. We assume there are no ties in the data. Let $t_{(1)} < t_{(2)} < \ldots < t_{(r)}$ be the ordered r distinct event times, so that $t_{(j)}$ is the j-th ordered death time. The set of individuals who are at risk at time $t_{(j)}$ will be denoted by $R(t_{(j)})$, so that $R(t_{(j)})$ is the set of individuals who are alive and uncensored at a time just prior to $t_{(j)}$. The quantity is called the risk set.

The relevant likelihood function for the proportional hazards model in equation (3.8) is given by

$$L(\beta) = \prod_{j=1}^{r} \frac{exp(\beta'x_{(j)})}{\Sigma_{l \in R(t_{(j)})}exp(\beta'x_l)'} \qquad (3.9)$$

in which $x_{(j)}$ is the vector of covariates for the individual who dies at the $j - th$ ordered death time, $t_{(j)}$. The summation in the denominator of this likelihood function is the sum of the values of exp $(\beta'x)$ over all individuals who are at risk at time $t_{(j)}$. The product is taken over the individuals for whom event times have been recorded. Individuals for whom the survival times are censored do not contribute to the numerator of the log-likelihood function, but enter into the summation over the risk sets at event times that occur before a censored time.

Suppose that the data consist of n observed survival times, denoted by $t_1, t_2, \ldots, t_n$, and that δ_i is censoring indicator which is zero if the i-th survival time $t_i, i = 1, 2, \ldots, n$, is right-censored, and unity otherwise. The likelihood function in equation (3.9) can be expressed in the form

$$\prod_{i=1}^{n} \left[\frac{exp(\beta'x_i)}{\Sigma_{l \in R(t_{(i)})}exp(\beta'x_l)} \right]^{\delta_i} \qquad . \qquad (3.10)$$

Where $R_{(i)}$ is the risk set at time t_i. In this study we propose to model the survival of teachers using covariates of age, gender, marital status, type of exit, teacher's working province, and job group using the following hazard model;

$$h_i(t) = h_0(t)exp(\beta_1 jobgroup_i + \beta_2 gender_i + \beta_3 age_i$$
$$+\beta_4 maritalstatus_i + \beta_5 exitname_i + \beta_6 Wprovince_i)$$

(3.11)

where, subscript i on an explanatory variable denotes the value of that variable for the ith individual. The $h_o(t)$ is the baseline hazard function , which is the function for an individual for whom all the variables are zero. In this study we examine the effect of the given explanatory variables above (which are the factors associated with exit from service) to the risk of leaving teaching. Fitting the above hazard model (equation 3.11), entails estimating the unknown β coefficients for the explanatory variables. The baseline hazard function, $h_0(t)$ also be estimated. These two components are estimated separately by first estimating the β coefficients using the method of maximum likelihood. In the treatment of ties , the equation (3.10) is modified. Breslow (1974) proposed the approximate likelihood in the presence of ties as;

$$\prod_{j=1}^{r} \frac{\exp(\beta' s_j)}{[\Sigma_{l \in R(t_{(j)})}\exp(\beta' x_l)]^{d_j}}$$

(3.12)

In this approximation, the $d_{(j)}$ deaths at time $t_{(j)}$ are considered to be distinct and to occur sequentially. The estimates of the explanatory variables in this study will be given along with their standard errors. These standard errors will be used to obtain approximate confidence intervals for the unknown β-parameters. The null hypothesis that $\beta=0$ can be tested by calculating the value of the statistic $\hat{\beta}_2/s.e.(\hat{\beta}_2)$. The observed value of this statistic is then compared to the percentage points of the standard normal distribution to obtain the corresponding P-value. This procedure is also called a Wald test. If there's no evidence to reject the hypothesis then that particular variable will be included in the model. However this statistic test

22

alone is not satisfactory, an alternative method for comparing different proportional hazards models will also be used. This is the statistic $-2log\hat{L}$. $\hat{L}$ is the maximized likelihood for a given model.

3.4.3 Comparing nested models

Suppose that two models are contemplated for a particular data set, model(1) and model(2), say, where model(1) contains a subset of the terms in model(2). Model(1) is then said to be parametrically nested within model(2). If the value of the maximized log-likelihood function for each model be denoted by $\hat{L}(1)$ and $\hat{L}(2)$, respectively. The two models can then be compared on the basis of the difference between the values of $-2log\hat{L}$ for each model. A large difference between $-2log\hat{L}(1)$ and $-2log\hat{L}(2)$ would lead to a conclusion that the q variates in model(2) that are additional to those in model(1) do improve the adequacy of the model. The amount by which the value of $-2log\hat{L}$ changes when terms are added to a model will depend on which terms have already been included. This gives the summary measure of agreement between the model and the data. $-2log\hat{L}$ will always be positive, and for a given data set, the smaller the value of $-2log\hat{L}$, the better the model. This statistic is only useful when making comparisons between models fitted to the same data.

3.4.4 Model selection

Identifying a set of explanatory variables that have the potential of being included in the linear component of a proportional hazards model is the initial step. This set will contain those variates and factors which have been recorded for each individual, but additionally terms corresponding to interactions between factors or between variates

been isolated, the combination of variables which are to be used in modeling the hazard function has to be determined. Practically, a hazard function will not depend on a unique combination of variables; instead, there are likely to be a number of equally good models, rather than a single 'best' model. Hence it is better to consider a wide range of possible models. An important principal in statistical modeling is that when a model contains an interaction term, the corresponding lower-order terms should also be included. The rule is, interactions should not be fitted unless the corresponding main effects are present. The model selection strategy depends to some extent on the purpose of the study. In this study, the explanatory variables are available and the aim of model selection is to determine which of them has an effect on the hazard function.

3.4.5   Variable selection procedure

First of all the explanatory variables are considered to be on an equal footing, and the aim is to identify subsets of variables upon which the hazard function depends. When the number of potential explanatory variables , including interactions, non-linear terms and so on, is not too large, it might be feasible to fit all possible combinations of terms, paying due regard to the hierarchic principle. Alternative nested models can be compared by examining the change in the value of $-2log\hat{L}$ on adding terms into a model or deleting terms from a model. If comparisons between a number of possible models, which need not necessarily be nested, can also be made based on the statistic;

$$AIC = -2log\hat{L} + \alpha q, \qquad \qquad \cdot(3.13)$$

where q is he number of unknown β-parameters in the model and α is a pre-determined constant. The value of α is usually taken to be between 2 and 6. This

statistic is known as Akaike's information criterion; the smaller the value of this statistic , the better the model. The value of AIC will tend to increase when unnecessary terms are added to the model. In this study we will identify subsets of variables upon which the hazard function depends, from our model of time to exit from teaching profession(equation 3.11).

Chapter 4

Results

4.1　Application of the survival function and Results

This study investigated the factors associated with teachers duration in service. This was done through modelling of these factors using survival analysis. In this case the beginning of the event is the date of hire of the teacher up to the end of the event when the teacher exits or leaves the teaching profession through resignation, death, dismissal, retirement at 50, mandatory retirement and transfer of service. This time has been calculated in years. For those who are still working at the time of data collection are censored. The software used is S plus 6.2. The first objective of this study was to find the length of the average teacher in service before exit. By using the Kaplan-Meier Estimator we are able to estimate the survival function of the teachers using the covariate type of exit. The covariate type of exit represents resignation, death, dismissal, retirement(at 50 and mandatory) ,and from the summary we are able to get the following results:

Table 4.1: Kaplan-Meier summary for types of exit.

Factor	Median (95% C.I)
Death	19.3 (18.4 - 20.2)
Dismissal	13.4 (11.3 - 15.9)
Resignation	13.0 (11.0 - 17.2)
Retire at 50	28.7 (27.3 - 29.3)
Retire Mandatory	30.3 (30.3 - 30.7)
Transfer service	16.9 (16.27 - 17.7)

The analysis shows that;

(a) Fifty percent of the teachers stay in the service for about 13 years before resigning.

(b) Fifty percent of the teachers who die stay in the service for about 19 years.

(c) Fifty percent of the teachers stay in the service for about 13 years before being dismissed by the TSC.

(d) Fifty percent of the teachers stay in the service for about 29 years before retiring at 50

(e) Fifty percent of the teachers stay in the service for about 30 years before retiring mandatory.

(f) And fifty percent of teachers stay in service for about 17 year before transferring their services.

Below are the Kaplan Meier of the survival curves for the covariate types of exit and when classified by gender and marital status.
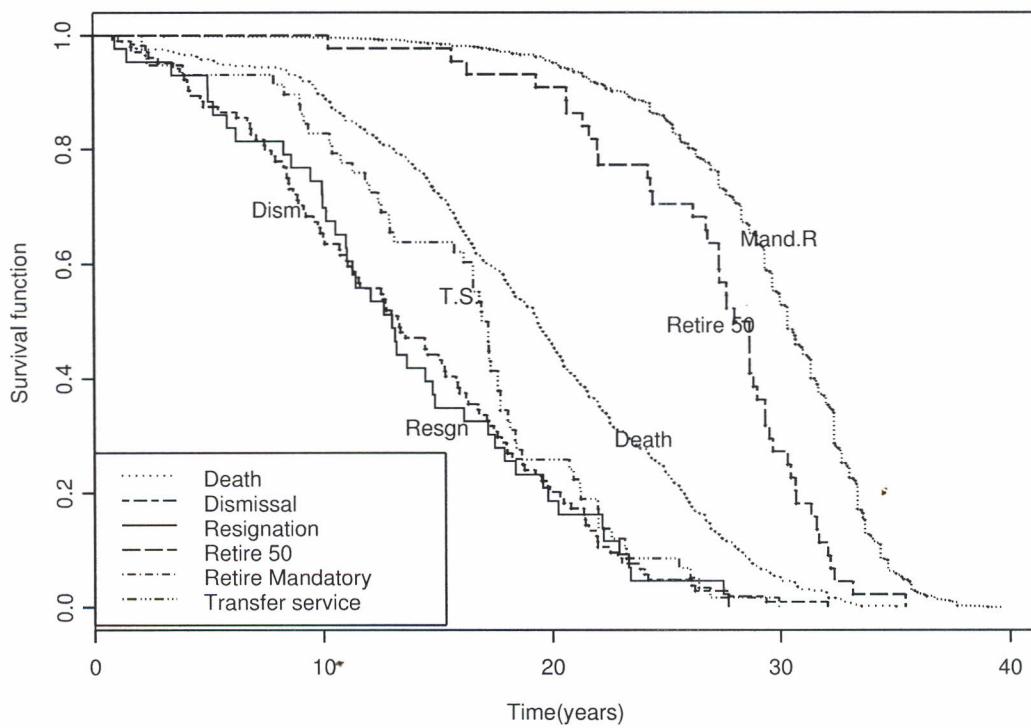
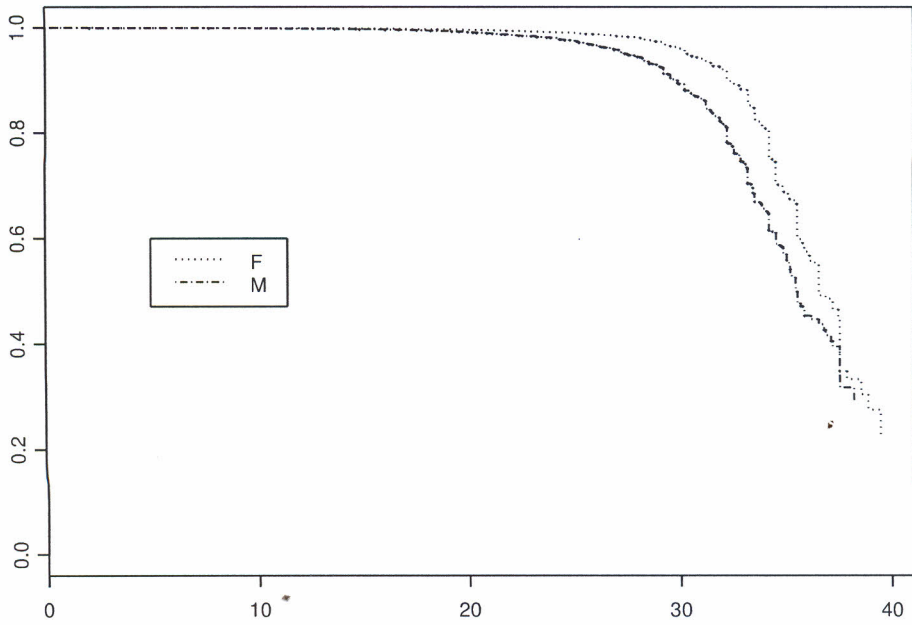Figure 4.1: Kaplan Meier estimate for survivor function of types of exit

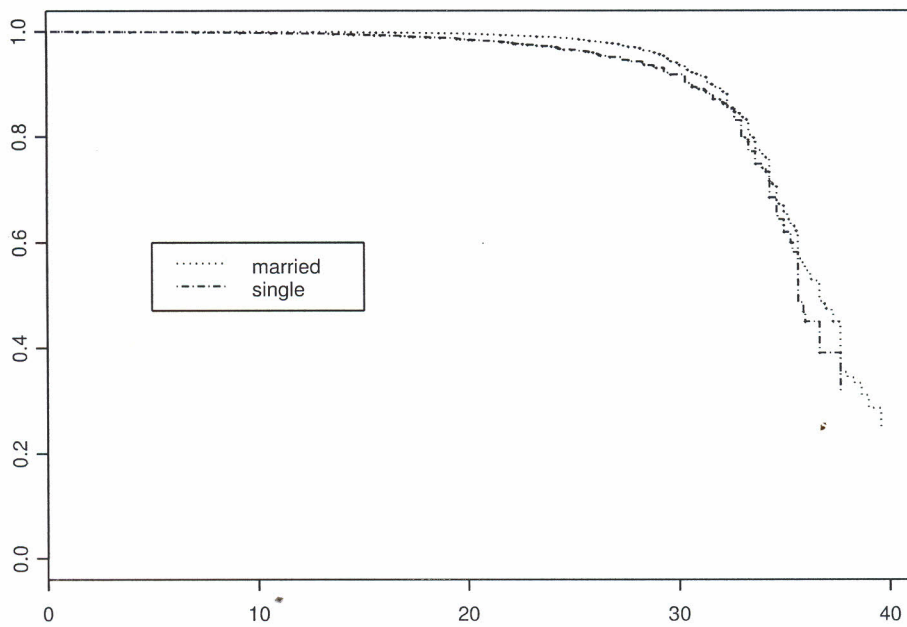Figure 4.2: Time to termination classified by gender

Figure 4.3: Time to termination classified by marital status

4.2   Comparing the survival curves

Table 4.2: Results of Log-rank test

Factor	N	Observed	Expected	$(O-E)\wedge 2/E$	$(O-E)\wedge 2/V$
Death	412	421	124.45	664.41	745.87
Dismissal	104	104	14.86	534.50	547.43
Resignation	43	43	5.73	242.49	244.98
Retire at 50	44	44	32.92	3.73	3.85
Retire Mandatory	1670	1670	2142.86	104.34	1414.33
Transfer service	58	58	10.18	224.57	228.30

Chisq= 1946 on 5 d.f, P value<0.0001 To find the differences in time to termination due to different types of exits, this study used the non-parametric method of log rank test. The null hypothesis, H_0: There's no difference in survival experience for the different types of exit vs. H_1: There are differences in survival experience for the different types of exit.

From the above results we find that the P-value< 0.0001 , we therefore reject the null hypothesis and accept the alternative one, that indeed there are differences in the survival experience due to different types of exit.

4.2.1   Discussion of the results

From Table 4.1 we see that the shortest length of time an average teacher can stay before exit is about 13 years through resignation and dismissal. This can be explained by the fact that teachers are looking for greener pastures in the case of resignation and many do so before teaching for long. In the case of dismissal, usually it's the TSC, the employer of the teachers that decides who is to dismissed

after breaching the TSC code of regulation. The TSC will dismiss any teacher found guilty irrespective of when the teacher was employed. Those who retire at 50 or at 55 stay longer in the classroom for about 29 and 30 years respectively. Those who transfer service on average stay in the classroom for about 17 years.

From the log-rank test (Table 4.2) we see that there are differences in time to termination by the different types of exit. The survival experience of the teachers varies with the type of exit, for example resignation, dismissal, transfer of service and death greatly reduces the survival experience of teachers as compared to retirement.

4.3 Factors associated with teachers' duration in service

4.3.1 Introduction

Cox regression model has been used to model the factors associated with teachers' duration in service. Here the variables job group,gender, age, marital status,type of exit and the teacher's working province which were collected with the date of hire and exit for each teacher, are included in the model of the survival of teachers given by equation 3.11. The subscript i on an explanatory variable denotes the value of that variable for the ith Individual. Job group,age, gender,marital status were assigned dummy variables as: marital status as S = single(includes divorced and widowed) and M = married; age as 1 = 40 years and below,2 = above 40 years; job groups as A = job groups B,D and E; G = job groups F and G; J = job groups H and J; L = job groups K and L; M = job groups M and N; P = job groups P,Q,and R; while gender was F for female and M for male.

4.3.2 Model building

In order to get the appropriate model, this study used the AIC procedure for variable selection, using the stepAIC function in the mass library of Splus. The stepAIC function removes the least helpful or least significant terms using the Akaike Information Criterion(AIC). The values of the likelihood ratio test was also used. The following complex model was fitted.

Start: AIC = 33723.67

$$Surv(Time.years, Censoring) \sim jobgroup + gender + age + maritalstatus$$
$$+ Exitname + Wprovince + gender * age * marital$$

$$(4.1)$$

We get the following results: The model selected is one without the variables

Table 4.3: Model selection results

step	AIC	$-2log\hat{L}$
1 Full model	33723.67	33673.67
2 -gender:age:marital	33632.25	33584.25
3-gender:age	33627.43	33581.43

gender:age:marital and gender:age which is given below

$$Surv(Time.years, Censoring) \sim jobgroup + gender + age + maritalstatus + Exitname+$$
$$Wprovince + gender : marital + age : marital$$

$$(4.2)$$

The model is then used for the analysis and the results are shown in table 4.4 and 4.5.

Table 4.4: Cox Ph output - final model

Factor	HR (95%c.i)	P value
Job group		
A	As reference	
G	1.125(0.88,1.43)	0.31
J	1.427(1.11,1.84)	6.0×10^{-3}
L	1.397(1.09,1.79)	9.0×10^{-3}
M	1.501(1.13,1.99)	4.9×10^{-3}
P	1.888(1.26,2.80)	2.0×10^{-3}
Gender		
female	As reference	
male	1.738(1.57 -1.92)	< 0.001
Age		
age(40 yrs and below)	As reference	
age(above 40 yrs)	0.083(0.07,0.10)	< 0.001
Marital status		
married	As reference	
single	4.756(3.45,6.58)	< 0.001
Type of exit		
Death	As reference	
Dismissal	1.144(0.88,1.48)	3.1×10^{-1}
Resignation	3.306(2.49,4.39)	< 0.001
Retire at 50	0.461(0.35,0.60)	< 0.001
Retire Mandatory	0.189(0.17,0.21)	< 0.001
Transfer service	2.926(2.31,3.71)	< 0.001
Working province		
Coast	As reference	
Eastern	0.692(0.50,0.96)	2.6×10^{-2}
Nairobi	1.43(1.11,1.85)	6.4×10^{-3}

Table 4.5: Cox Ph output - final model

Factor	HR (95%c.i)	P value
North Eastern	0.462(0.19,1.14)	9.5×10^{-2}
Nyanza	0.997(0.74,1.35)	9.8×10^{-1}
R.Valley	1.021(0.83,1.25)	8.4×10^{-1}
Western	1.512 (1.25,1.83)	$P < 0.001$
Central	1.00(0.82,1.22)	9.9×10^{-1}
gender:marital	0.475(0.37,0.61)	$P < 0.001$
age:marital	0.457(0.33,0.63)	$P < 0.001$

4.3.3 Discussion of the Results

From the Cox Ph output above we get the following:

Job group

As indicated above the reference job group is here denoted by A which is the combined job groups of individual job groups B,D and E. Job group G, a combined job groups of individuals job groups F and G has a hazard ratio (HR) of 1.125(P=0.34) indicating that teachers in job groups F and G, increases the risk to termination by 1.13 fold relative to those in job groups B,D and E. The P value of 0.34 indicates that there's no significant difference in risk between teachers in lower job groups B,D and E and those in job groups F and G. This can be explained by the fact that teachers in these job groups are designated as P1, P2 and P3 teachers who rarely leave the service but instead opt to further their studies.

Job groups J denoting individual job groups H and J has a hazard ratio 1.427(P= 0.006), meaning that teachers in job groups H and J increases the risk to termination by 1.427 fold relative to those in job groups B,D and E and there's significant difference in risk between these teachers. Job groups L denoting individual job groups K and L, M denoting individual job groups M and N, P for individual job groups Q,R

35

and P have hazard ratios of 1.397 (0.009),1.501(0.0049) and 1.89(0.002) respectively; indicating that teachers in these job groups increase the risk to termination by 1.397, 1.50, and 1.89 fold respectively relative to those in job groups B, D and E . From the P values we can conclude that they are significantly different in risk from the reference lower job groups B,D and E. This can be explained by the fact that teachers in these job groups are more qualified and may therefore find it easier to get better jobs. In addition most teachers in these higher job groups are already in the retirement bracket and hence their higher risk to termination.

Gender

Female teachers is the reference as indicated above. Gender (males) has a HR of 1.738 ($P < 0.001$) meaning that a male teacher increases the risk to termination by 1.738 fold relative to a female teacher and there's a significant difference in risk between the male and female teacher. This may be explained by the fact that more male teachers may be more aggressive in looking for greener pastures or more of them are being dismissed. We can conclude that gender is significantly associated with the risk to termination since the confidence interval of HR does not contain unity.

Age

Age 1 representing those aged 40 years and below has been used as the reference age. Therefore age 2 representing those aged above 40 years has a hazard ratio of 0.083($P < 0.001$). We conclude that being a teacher aged 40 years and above reduces the risk to termination by 91.7% and hence, have longer expected survival time than those aged 40 years and below. The confidence interval of the hazard ratio does not include unity hence there's significant difference in risk for teachers who are aged above 40 years to those who are 40 years and below; age is significantly associated with the risk to termination. This may be explained by the fact that younger teachers are more aggressive and experimental and may easily find their

way to greener pastures as compared to the older teachers who have accumulated more specific capital that is knowledge specific to occupation and that which is non transferable (Kirby and Grissmer, 1993).

Marital status

Marital status has a hazard ratio of 4.756 ($P < 0.001$). This is the hazard ratio of single teachers relative to married teachers. Therefore being a single teacher increases the risk to termination by 4.756 fold relative to a married teacher, hence have shorter expected survival time than the married teachers. It's P value is significant indicating that there's significant difference in risk for single teachers to those who are married. Marital status is significantly associated with the risk to termination.

Types of exit

Death has been used as reference. Dismissal, resignation and transfer of service have HR of 1.14 (P=0.31), 3.31 ($P < 0.001$) and 2.93 ($P < 0.001$) respectively relative to death. Dismissal of teachers increases the risk to termination by 1.14 fold relative to death, the p value and the confidence interval of the hazard ratio indicates that there's no significant difference in risk by dismissal to that of death. Resignation increases the risk to termination by 3.31 fold relative to death. The confidence interval of the hazard ratio does not include unity so we conclude that there's significant difference between resignation and death. Transfer of service increases the risk to termination by 2.93 relative to death and the two are significantly different.

Retirement at 50 and mandatory retirement have a HR of 0.461($P < 0.001$) and 0.189($P < 0.001$) respectively relative to death. Hence retirement at 50 reduces the risk to termination by 54% while mandatory retirement by 81%. From their P values there's indication they are significantly different from death.

From the results resignation, transfer of service and dismissal reduces the survival

rates of teachers. This can be explained by the fact that teachers terminate service with the TSC for greener pastures while others exit the service through dismissal due to breaching the TSC codes of regulation. Retirement does not reduce the survival rates of teachers relative to death since one has to work for long before retirement.

Working province

Coast has been used as the reference as shown above. Eastern has a hazard ratio of 0.692 (P=0.026) indicating that working in this province decreases the risk to termination by 31% relative to working in Coast Province. Working in Nairobi increases the risk to termination by 1.43 fold relative to working in Coast Province and from the p value we conclude that there's significant difference in risk of a teacher working in Nairobi to that working in Coast Province. Working provinces North Eastern and Nyanza have HR of 0.462 (P=0.095) and 0.997 (P=0.98), hence they reduce the risk to termination by 54% and 0.003% respectively. From the p value of Nyanza we can conclude that there's no significant difference in risk of a teacher working in Nyanza to the one in Coast. Rift Valley, Western and Central provinces increases the risk to termination by 1.021 (P=0.84), 1.512 ($P < 0.001$) and 1.00 (P=0.99) respectively. However from the p value there's no significant difference in risk of working in Rift Valley and Central relative to Coast province. Teachers in certain provinces like Nairobi which is urban are more likely to get better alternative jobs compared to those in provinces like North Eastern and Eastern.

Interactions

The interactions between gender:marital status (single male) and age:marital status (single teacher aged above 40years) have a HR of 0.475 ($P < 0.001$)and 0.457 ($P < 0.001$) and parameter estimates of -0.745 and -0.784 respectively. A male teacher who is single has a 2.258 fold of the risk to termination relative to a male married teacher, while a female single teacher has a risk of 4.754 fold higher than

a female married teacher. A married teacher above 40 years has a HR of 0.0829 relative to a married teacher aged below 40 years hence being a married teacher aged above 40 years reduces the risk to termination by 91.7% relative to a married teacher aged below 40 years.

We therefore conclude that the correlates to termination are job group, gender, age, marital status, type of exit, working province, gender:marital status and age:marital status.

Chapter 5

Conclusions and Recommendations

This study was concerned with modelling factors associated with teachers duration in service using survival analysis and especially the cox regression hazard model which allows the assessment of the impact of multiple covariates in the same model. The study specifically sought to find out how long an average teacher can stay in the classroom before exit, whether there are differences in time to termination by types of exit and to determine the correlates of termination of service. The study established that an average teacher can stay in the classroom for about 13years before resigning or being dismissed, 17 years before transferring service,19 years before death, 29 and 30 years before retiring at 50 or mandatory retirement. In view of these findings resignation then transfer of service,dismissal,and death in that order greatly reduces the survival experience of teachers . This means that teachers exit from service before staying for long especially through resignation, dismissal and transfer of service. The government will continue to replace such teachers with other newly trained ones implying more costs. There's need to address the issue of retaining the already recruited and trained teachers to save the government's bill on education. Transfer of service needs to be addressed. Dismissal of teachers could

could also be checked by guiding and counseling the teachers concerned and giving them more room for change before their dismissal. From the results of the final model we can conclude that job group, gender, age , marital status, type of exit, working province, interaction of gender and marital status and that of age and marital status are the factors associated with teachers' termination from service in this study. The risk to termination increases with higher job groups. The Ministry of Education could also organize regular seminars and inservice training for newly recruited teachers to help them develop skills critical for success in the profession. Successful teachers are more likely to stay in the profession. Further work can be done in the same area by using a different approach like parametric models.

Bibliography

[1] Bobbit,S.A., Faupel,E.,and Burns,S. (1991).*Characteristics of stayers, movers, and leavers. Results from the teacher follow-up surveys, (1988-1989)*(NCES publication No. 91-128). Washington, DC: US Department of Education, office of Educational Research and Improvement.

[2] Boe, E.E.,Bobbit,S.A.,and Cook,L.H. (1993). *Whither did st thou go: Retention, reassignment, migration, and attrition of special and general education teachers in national perspective.* Paper presented at the Annual convention of the Council for Exceptional children, San Antonio,TX.

[3] Boe,E.E., Cook,L.H., Bobbit,S.A.,and Weber,A.L. (1995). *Retention, transfer and attrition of special and general education teachers in national perspective.* Paper presented at the National Dissemination Forum on Issues Relating to Special Education Teacher Satisfaction, Retention and Attrition,Washington, DC.

[4] Bloland,P.A., and Selby,T.J. (1980). Factors associated with career change among secondary school teachers: A review of the literature. *Educational Research Quarterly,* 5(3), 13-24.

[5] Charter,W.W.,Jr., (1970,Jan). Some factors affecting survival in school districts. *American Educational Research Journal, Vol.7, No.1, pp. 1-27.*

[6] Collet,D.(1994). *Modelling survival data in medical research.*London:Chapman and Hall.(1-106).

[7] Grissmer,D.W., and Kirby,S.N. (1987). *Teacher attrition: The uphill climb to staff the nation's schools.*Santa Monica:Rand.

[8] Heynes,B (1988,April). Educational defectors: A first look at teacher attrition in the NLS-72. *Educational Researcher,* 17(3), 24-32.

[9] Kenya National Bureau of Statistics and Ministry of planning and National development (2007). *Economic survey 2007,* Government printer, Nairobi. 41-51.

[10] Kenya National Bureau of Statistics and Ministry of planning and National development (2007). *Statistical abstract 2007,* Government printer, Nairobi.

[11] Keily,M.K., and Martin,N.C. (2005) Survival Analysis in Family Research. *Journal of Family Psychology,* Vol.19, No.1, 142-156.

[12] Kirby,S.N., and Grissmer, D.W. (1993). *Teacher attrition:· theory, evidence and suggested policy options.* Paper presented at the world bank/ Harvard institute for International Development seminar on "Policies affecting learning outcomes through impacts on teachers" held at Cambridge, MA.

[13] Mark,J.H., and Anderson,B.D. (1985). Teacher survival rates in st. Louis, 1969-1982. *American Educational Research Journal,* 22, 413-421.

[14] Mark, J.H.,and Anderson,B.D. (1978). Teacher survival rates - A current look.*American Educational Research Journal,* 15, 379-383.

[15] Murnane,R.J., Singer,J.D., Willet, J.B., Kemple,J.J., and Olsen, R.J. (1991).*Who will teach? Policies that matter.* Cambridge: Harvard university press

[16] Murnane,R., Singer,J.D.,and Willet, J.B. (1989). The influences of salaries and "opportunity costs" on teachers'career choices: Evidence from North Carolina. *Harvard Educational Review,* 59, 325-346.

[17] Murnane,R., Singer,J.D.,and Willet, J.B. (1988). The career paths of teachers: Implications for teacher supply and methodological lessons for research.*Educational Researcher,*17(6), 22-30.

[18] Schlecty,P.C., and Vance,V.S. (1981). Do academically able teachers leave education? The North Carolina case,*phi Delta Kappan,* 63, 106-112.

[19] Singer,J.D and Willet, J.B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence.*(pp 300-310) New York: Oxford university press.

[20] Singer,J.D., and Willet, J.B. (1988). Detecting involuntary layoffs in teacher survival data: The year of leaving dangerously. *Educational Evaluation and policy Analysis,* 10, 212-224.

[21] Theobald,N.(1990). An examination of the influence of personal, professional, and school characteristics on public school teacher retention.*Economics of Educational Review,* 9, 241-250.

[22] Nairobi EFA coordinators' consultation (2006, July). *Kenya country status report.*

[23] Whitener,J.E.(1965). *An actuarial approach to teacher turnover.* Unpublished doctoral dissertation. St.Louis, MO: Washington university.

[24] Willet, J.B., and Singer,J.D. (1991,Winter). From whether to when: New Methods of Studying Student Dropout and Teacher Attrition. *Review of Educational Research,* Vol. 61, No.4, pp. 407- 450.

Appendix A

Splus Script

```
options(contrasts=c("contr.treatment","contr.poly"))
table(SDF4$Exitname)
mat< -SDF4[SDF4$Exitname!="WK",]
dim(mat) mat< -as.data.frame(mat)
table(mat$Exitname)
summary(mat)
fit< -survfit(formula = Surv(Time.years., CENSORING, type = "right") ~Exitname,
data = mat, na.action = na.exclude, se.fit = T, type
= "kaplan-meier", error = "greenwood", conf.type = "log", conf.lower
= "usual")
summary(fit)
plot(fit,lty=c(2,6,1,4,7,5),xlab="Time(years)",ylab="Survival function") legend(0,0.27,c("Death","Dismis
"Retire 50","Retire Mandatory", "Transfer service"),lty=c(2,6,1,4,3,5))
text(15,.32,'Resgn')
text(8,.7,'Dism')
text(16,.55,'T.S.')
```

```
survdiff(Surv(Time.years.,CENSORING)~ mz, data=SDF7, rho=0)
survfit(mz)
plot(survfit(mz),lty=2:3)
legend(5,0.6,c("married", "single"),lty=2:3)


ml< -coxph(formula = Surv(Time.years., CENSORING, type = "right")~jobgroup
+ gender + age1 + marital1 + WProvince + strata(gender), data = SDF7, na.action =
na.exclude, method = "efron", robust = F)
survfit(ml)
plot(survfit(ml),lty=2:3)
legend(5,0.6,c("F", "M"),lty=2:3)


options(contrasts=c("contr.treatment","contr.poly"))
Z2< -coxph(formula = Surv(Time.years., CENSORING, type = "right") ~jobgroup1
+ gender + age1 + marital1 + Exitname+ WProvince + gender*age1*marital1,
data = SDF7, na.action = na.exclude, method = "efron", robust = F)
```
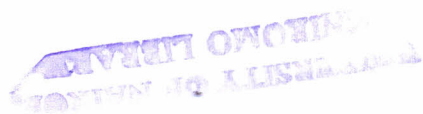
Z2< − stepAIC(coxph(formula = Surv(Time.years., CENSORING, type = "right") ~ jobgroup1 + gender + age1 + marital1 + Exitname + WProvince + gender * age1 * marital1, data = SDF7, na.action = na.exclude, method = "efron", robust = F), direction = "both")
Z2$anova

Zb< −coxph(formula = Surv(Time.years., CENSORING, type = "right") ~jobgroup1 + gender + age1 + marital1 + Exitname+ WProvince + gender:marital1+age1:marital1, data = SDF7, na.action = na.exclude, method = "efron", robust = F)
summary(Zb)