

A Semi-Parametric Estimation of Time to Exit from Service of Teachers in Kenya

Jerita Jemimah Mwambi[1]*, Leonard Kiti Alii[2]

1. School of Humanities and Social Sciences, Pwani University, P.O Box 195, Kilifi 80108, Kenya
2. School of Pure and Applied Sciences, Pwani University, P.O.Box 195, Kilifi 80108, Kenya

*E-mail of Corresponding Author: jerrymimah2012@gmail.com

Abstract

There has been increased enrolment in both primary and secondary schools in the country 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 those already in the service. In order to achieve this, there is need to investigate the factors associated with teachers' exit from service. This study presents the semi-parametric Cox proportional hazard model to determine a combination of potential explanatory variables that are associated with survival time in service for teachers employed by the Teachers Service Commission. The study used data from the Teachers Service Commission of 278,063 teachers working in public schools and post primary institutions in Kenya up to October 2014 and 28,403 who have left the Teachers Service Commission. Results show that male teachers have a higher hazard ratio (1.23) compared to their female counterparts hence indicating that male teachers are more likely to leave teaching, older teachers aged 35 years and above reduces the risk to termination by 96.9% relative to those aged 35 years and below. Teachers working in North Eastern parts of the country have lower survival rates hence are more likely to exit than those in other provinces.

Keywords: The Survival function, Cox Proportional Hazard model, Teachers' Survival Time in Service, Time to exit

1. Introduction

Education is one of the strong social pillars of national development in a country (Kenya Vision, 2030). One of the major objectives of the government of Kenya is to provide quality basic education and training to every citizen; which is fundamental to the success of the government's overall development strategy (Education International, 2007). Education has a very big role to play in the realisation of the vision 2030 and the achievement of the Education for All (EFA) targets in Kenya and research has shown that a qualified teacher is a very important factor in the achievement and performance of students (Education International, 2007; International Task force on teachers, EFA, 2010). Since the Kenya government implemented the free primary education in 2003, enrolment has been increasing at a very high rate (UNESCO, 2014, Kenya Sessional paper, 2005). According to the Economic survey (2012), total enrolment of pupils in standard one to eight in public primary schools was 9.86 million in 2011 up from 9.38 million in 2010; while in secondary schools enrolment had increased from 1.65 million in 2010 to 1.77 million in 2011 in secondary schools. Moreover, the pupil: teacher ratio in primary schools went up from 54:1 in 2010 to 57:1 in 2011(Economic Survey, 2012). This adversely affects the learning process as interaction between an individual pupil and the teacher is very limited (UNESCO, 2005). The quality of learning is adversely affected by such big ratios (GCE/Action Aid 2005a: 21-34). Adequate number of teachers is required to ensure every school going child gets the best of the education that can be provided (UNESCO, 2013/4).

However, one of the major challenges facing the country, like in many sub-Saharan countries, is teacher attrition (Education International 2007; International task force for EFA, 2010). A survey on teacher supply, recruitment and retention in six Anglophone countries, Kenya included, reported an average rate of attrition of 4% and this is mainly attributed to retirement, resignation, death and dismissal (Education International ,2007), However, due to the large proportion of voluntary resignation attrition rate is likely to rise (International task force for EFA 2010). According to data-based evidence of impact of HIV/AIDS on teachers, Kenya, Tanzania, and Zambia would each lose 600 teachers to AIDS alone in 2005(UNESCO, 2006a: 87-88).

In Kenya the government is the main employer of teachers through the Teachers Service Commission (TSC), which was established by an act of parliament in 1967. However, in order to contain the wage bill strategy the TSC has only been replacing teachers who exit the service through natural attrition (Education International, 2007). This challenge needs to be addressed for the country to enhance equitable and quality education; and to develop industrially and technologically by 2030.

This study explored the factors associated with termination of teachers' service by using survival analysis methods and in particular the Cox proportional hazard model (Cox, 1972). This model is appropriate since it deals with an outcome variable of interest which is time to exit in an employment context. Survival analysis provides straight forward easily application tools for appropriately analyzing the timing of educational transitions that is, time from entry to exit from career. The start or end points for measuring times are chosen according to the context and may not represent the entire life of an individual. The time to the event of interest is called failure time or survival time. Study of time to the event of interest has been widely applied in many areas of social sciences and biostatistics. The survival time of an individual is said to be censored when the end point of interest has not been observed for that individual, or an individual is lost, migrates or withdraws from the study during the follow up period. In this study the event may not have occurred for some teachers at the end of the follow-up period, these are teachers who do not exit the service and will possess censored event times. It is not known when or whether the teachers will experience the event; all we know is that by the end of the data collection, the event had not yet occurred.

Various past studies have been carried out to unravel the causes and effects of teachers' exit from service (Charters 1970; Mark and Anderson 1978; Murnane 1981). Past studies reveal that personal characteristics like level of education, salary, experience, marital status, age, gender, subject speciality and working place and conditions contribute to teachers' exit from service. Bowman and Dowling (2008) carried out a comprehensive meta-analysis of all quantitative studies on teacher attrition and retention. They sought to understand why attrition occurs or what factors contribute to it. In their analysis they found that personal characteristics, working environment and resources available influence teacher attrition and retention.

Research has shown that age of the teacher influences the decision of whether to leave the profession or not; and that the relationship between the teacher's age and turnover depicts a U-shaped curve (Ingersoll, 2001). Younger teachers tend to have high rates of exit while those in the mid-career tend to stabilise, then the rate rises during the retirement period (Ingersoll, 2001; Singer and Willet, 1988).

Crandell and Howell (2009) carried out a study on the three kinds of teacher attrition: school transfer, exit and retirement by employing a multinomial logistic model for competing risks. They used the cross sectional data from Schools and Staffing Survey (SASS) and the Teacher Follow up Survey (TFS) data. They sought to find out whether salary increases are predicted to reduce teacher attrition and whether there are differential effects on the various types of attrition. They also focused on factors like school level, urbanicity, teacher education, student behavior, administrative supportiveness, teacher's individual characteristics, and years of experience, race, family and marriage variables, attainment of a master's degree and beyond. One of their major finding is that salary is a significant predictor of transfer and retirement attrition but not exit attrition. Their study finds salary not to be significant at a mere 10% significant level for leavers, indicating that salary may have a little effect on teachers' decisions to leave the profession, though a popular policy recommendation is to increase teachers' salary.

Adams (1996) used the Cox proportional hazard model to examine the career paths of 2,327 elementary teachers hired by a large Texas School district. He examined the influence of the variables age, sex, and ethnicity and certification route to the survival of teachers. His results indicated that women had a hazard of approximately 37% of leaving the district compared to men; while younger teacher below 40years were more likely to leave the district than older teachers.

Imazeki (2005) used data from Wisconsin to study teacher labour mobility within and out of the teaching profession and employed a competing risk duration model to examine the influence of teacher and district characteristics like various measures on salary, time spent in a district and reasons for leaving. She noted that most studies on teacher attrition, estimate probability of attrition by using the traditional binary logit or probit model which are limiting since they fail to distinguish the effect of variables on exits and transfers separately. The study found that transfers responded most strongly when district salaries were increased relative to nearby districts. Salary increases for more experienced teachers reduced exit attrition among newer female teachers; hence increasing salaries for beginning teachers can help reduce high attrition rates in districts that lose many teachers to other districts like rural districts.

In African developing countries some attrition is encouraged by conditions within the teaching such as poor living conditions of teachers. Teachers are not housed within schools and find it difficult to get affordable accommodation and the few that are housed live in deplorable conditions (International Task force on Teachers for EFA ,2010). Schools located in remote rural areas are usually understaffed as they lack proper teacher housing and because of poor transport, there is lack of services like hospitals, banks, libraries, shopping and leisure facilities (VSO, 2002). The study also found that teachers in such schools are rarely visited by regional and provisional authorities and do not receive prompt communications from the offices. Teaching and learning materials are slow to reach such schools. Teacher attrition tends to be higher in such rural schools and usually end up with unqualified teachers (kadzamira, 2006).

Most of the studies on teacher attrition estimate probability of attrition by using the traditional binary logit or probit model. These are limiting since they focus on the static nature of exit; that is, whether a teacher exits or not instead of both whether or not a teacher exits and when.

Substantial research on survival rates of teachers and attrition has been done worldwide. However very little research effort in Kenya has been directed into finding which teacher leaves, when and why in public schools and institutions of Kenya by survival analysis methods.

2.0 Materials and Methods

The data used in this study was obtained from the Teachers' Service Commission consisting of 278,063 teachers employed by the government working in public primary schools and post primary institutions in Kenya up to October 2014 and 28,403 teachers who have exited through death, retirement (mandatory, retire at 50, retire due to public interest, on medical grounds, voluntary early retirement), dismissal, transfer of service, resignation, contract expiry. This is after teachers with missing information were removed from the data. This data was collected with permission from the Ethics review Committee, Pwani University accredited by the National Commission for Science, Technology and Innovation, KENYA and the Teachers' Service Commission. The variables job group, age, gender, type of exit, and the province of the teacher's working station and. the date of hire and exit (if applicable) were availed for each individual teachers. The salary of the teacher was calculated using the job group from the information availed by the TSC. 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) 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 were not censored and indicated by (1). The analysis was done using R-3.1.3 software (R development core team, 2015).

2.1 The Survivor function

It is also known as the Survival probability and it is the probability of surviving beyond sometime t or the probability that an event of interest has not yet happened by time t and it is denoted by S (t) ; such that

$$S(t) = P(T > t)$$

We estimated the survival function by using the Kaplan- Meier (KM) estimator. It is also known as the product-limit estimator (Kaplan and Meier, 1958). The KM survival curve is a plot of the KM survival probability against time, which gives a good summary of the data and can also be used to estimate the median survival times and other percentile measures. The steeper the survival curves the lower the survival rate. The KM survival curves for a categorical variable can also be used as a graphical test for the proportional assumption in that, when the curves cross, then the proportional assumption has been violated.

2.2 The Hazard function

The hazard function or rate is a risk function (kalbfleisch & Prentice, 2002; Collet, 2003). It gives the probability that an individual fails at time t, given that he/she has survived up to time t. The hazard function is given by;

$$\lambda(t,x) = \lim_{\Delta t \to 0} P(t \le T < t + \Delta t / T \ge t) / \Delta t \quad (1)$$

The Cumulative hazard function is given by;

$$H(t) = \int_0^x \lambda(u)du = -\ln[S(x)]$$

For continuous lifetimes;

$$S(t) = [\exp(-H(x)] = \exp[-\int_0^x \lambda(u)du]$$

2.2 *Comparison of survival curves*
We used Log rank test to determine if groups of teachers according to gender, age, job groups and the province of the working station have different survival experiences in time to termination of service (See Table 3). The non-parametric Log-rank test statistic is more appropriate than the K-M survival curves. The null hypothesis is H_0: There is no difference in survival experience across the groups versus the alternative hypothesis, H_1: There are differences in the survival experience across the groups.

The log rank test is given by:

$$\chi_p^2 = \sum_{g=1}^{k} \frac{(O-e)^2}{e} \quad (2)$$

Where O is the observed and e, the expected number of death in the gth group and the test has a Chi-squared distribution on (k-1) degrees of freedom.

2.3 The Cox proportional hazard (PH) model
Since we are dealing with time to event data and some teachers are censored we cannot use multiple linear regression or logistic regression. One method which is commonly used for such kind of data is the popular Cox PH model.
We applied the Cox (PH) model to estimate the time to exit for teachers along with their explanatory variables already mentioned above. The Cox PH models the hazard function, $\lambda_i(t)$ for the ith individual at time t,

$$\lambda_i(t) = \exp(\beta_1 z_{1i} + \beta_2 z_{2i} + ----+ \beta_p z_{pi})\lambda_0(t) \quad (3)$$
$$= \lambda_0 \exp(\beta')$$

$\beta_{1'}..........\beta_p$ are the regression coefficients and z_i is the explanatory or predictor variable.

$\lambda_0(t)$ is the baseline hazard function corresponding to $Z = 0$.

The Cox PH model is widely applied in failure data analysis since it makes no distribution assumption for λ_0 the baseline hazard function. The parametric form is only assumed for the covariate effect, therefore the name Semi-parametric model.
The corresponding Survivor function is given by:

$$S(t) = S_0(t)^{\exp(Z\beta')} \qquad (4)$$

Where

$$S_0(t) = \exp[-\lambda_o(u)du]$$

This is the survival function of the sub population with covariate Z.

Now if Z_1 and Z_o are values of covariates for two individuals, then ratio of their hazard rates is

$$\frac{\lambda(t/z_1)}{\lambda(t/z_o)} = \frac{\lambda_0(t)\exp(z_1^{'}\beta)}{\lambda_0(t)\exp(z_0^{'}\beta)} = \exp(z_1 - z_0)'\beta,$$

Which is a constant for all t and it is known as the hazard ratio or relative risk, hence the Cox model is usually referred to as a proportional hazards model.

2.4 Estimation of the β Parameters

Fitting the proportional hazards model given above in equation (2) to an observed data requires the estimation of the unknown coefficients of the explanatory variables $Z_1, Z_2, ..., Z_P$, in linear component of the model, that is $\beta_1, \beta_2, \ldots \beta_p$. The baseline hazard function λ_0 (t) may also be estimated. One way of estimating the $\beta^{'s}$ separately is by use of maximum likelihood method. Cox (1972) proposed a partial likelihood technique for estimating the regression parameters while λ_0 is taken as a nuisance parameter. By maximising the partial likelihood function given below (equation 5), consistent and efficient estimators of the regression parameters can be obtained.

$$L(\beta) = \prod_{j=1}^{r} \frac{\exp(\beta^{'} z_{(j)})}{\sum_{q \in R(t_{(j)})} \exp(\beta' z_i)'} \qquad (5)$$

This partial likelihood is obtained as shown below. Let the ordered survival times for n individuals be: $t_1, t_2 \ldots, t_n$. Suppose $t_{(1)} < t_{(2)} < \ldots < t_{(r)}$ are ordered r death times, assuming there are no ties. $R(t_j)$ is the risk set just prior to time t_j. The probability that the ith individual fails at time t_j conditional on t_j being one of the event times is given by;

$$P[\,ith\ individual\ \ fails\ at\ t_j / t_j\ being\ an\ event\ time]$$

$$= \frac{P[an\ individual\ with\ covariate\ z_j\ \ fails\ at\ t_j]}{P[one\ event\ at\ time\ t_j]}$$

$$= \frac{P\,[an\ individual\ with\ \text{cov}ariate\ z_j\ \ fails\ \ at\ t_j]}{P\,[one\ event\ at\ time\ t_j]}$$

$$= \frac{P\,[an\ individual\ with\ \text{cov}ariate\ z_j\ \ fails\ \ at\ t_j]}{\sum_{q\in R(t_j)} P\,[individual\ q\ \ fails\ \ at\ time\ t_j]}$$

$$= \frac{P\,[an\ individual\ with\ \text{cov}ariate\ z_j\ \ fails\ \ in\ \{t_j,(t_j+\Delta t)\}/\Delta t}{\sum_{q\in R(t_j)} P\,[individual\ q\ \ fails\ in\ time\ \{t_j,(t_j+\Delta t)\}/\Delta t]}$$

$$= \frac{\lim_{\Delta t\to 0} P\,[an\ individual\ with\ \text{cov}ariate\ z_j\ \ fails\ \ at\ \{t_j,(t_j+\Delta t)\}/\Delta t}{\lim_{\Delta t\to 0} \sum_{q\in R(t_j)} P\,[individual\ q\ \ fails\ at\ time\ \{t_j,(t_j+\Delta t)\}/\Delta t]}$$

$$= \frac{\lambda(t_j)}{\sum_{q\in R(t_j)} \lambda_q(t)}$$

$$= \frac{\lambda_0(t\ )\exp(\beta' z_{ij})}{\sum_{q\in R(t_j)} \lambda_0(t)\exp(\beta' z_q t_j)}$$

$$= \frac{\exp(\beta' z_{ij})}{\sum_{q\in R(t_j)} \exp(\beta' z_i t_j)}$$

The partial likelihood function for the Cox PH model is then given by

$$\ell(\beta) = \frac{\prod_{j=1}^{r} \exp(\beta' z_i t_j)}{\sum_{q\in R(t_j)} \exp(\beta' z_q t_j)} \qquad (6)$$

Where $z_i(t_j)$ is the vector of covariate values for individual i who fails at time t_j.
This likelihood function can be expressed as

$$\ell(\beta) = \prod_{i=1}^{n} \left[\frac{\exp(\beta' z_i t_j)}{\sum_{q\in R(t_j)} \exp(\beta' z_q t_j)} \right]^{\delta_i} \qquad (7)$$

where δ_j is the censoring indicator which equals to zero if the survival time is censored and one otherwise.
The partial likelihood function can only be used when there are no ties in the data set,
To accommodate ties the Breslow method (Breslow, 1974) is used;

$$\prod_{j=1}^{r} \frac{\exp(\beta' s_j)}{\left[\sum_{q \varepsilon R(t_j)} \exp(\beta' z_q) \right]^{\delta_j}} \qquad (8)$$

where s_j is the vector of sum of each of p covariates for those individuals failing at t_j. Most statistical software can easily handle ties by breslow method. In our study we use the Breslow method in R since there are ties in our data set, that is, where two or more teachers have the same survival time.

2.5 Statistical Analysis

Statistical analysis was done using the statistical software R 3.1.3 (R development Core Team, 2015) as already stated above. Univariate analysis with the Cox PH model for each explanatory variable was conducted. P values < 0.05 were considered to be significant and were used for multivariate analysis using the Cox PH model to build the factors associated with the survival time of teachers in service. The step function in R for backward selection of the variables was used to get the appropriate model.

Graphical assessment for the proportional assumption of the model was done using the Kaplan- Meier plots of categorical predictors such as age, gender and job group. The proportional assumption is violated when curves of the groups cross.

3.0 Results

Non-parametric tests were carried out using the Kaplan-Meier estimator (Kaplan & Meier, 1958) and log rank test. To investigate the survival experience of teachers classified by gender and age we employed the K-M estimator. We had a total of 166, 824 male teachers (coded as M) and 139, 642 female teachers (coded as F). We also had a total of 57,471 teachers who are less than or equal to 35 years (coded as 0), while those who are above 35 years were 248,995 (coded as 1). The explanatory variable for job groups had eleven levels initially from job group F to R but we reduced them to four levels as: Job group A (for job groups F, G and H), B (for job groups J, K and L), C (for M and N) and D (for job groups P, Q and R). Types of exit were also reduced to five levels. Individuals who had exits with unknown reasons were deleted from the data. The covariate province representing the region where the teacher is working included all the eight provinces or regions in Kenya. We decided to use the provinces instead of counties as they would result into too many levels.

3.1 The survival function and Log rank test.

The KM plot summary gives the median survival times for female (38.2,) and male teachers (36.4) in years as shown in Table 1.The K-M survivor curve for gender is given in figure 1.

The K-M survivor function for age gave the median survival times of teachers aged 35 years and less as 12.3 years while those older than 35 years as 37.0 years (see Table 2). Figure 2 shows the K-M curve for covariate age.

We employed the non-parametric Log- rank test statistic (equation 2) to determine if groups of teachers classified by covariates age, gender, province and job groups have different survival experiences. P values for covariates age, gender, and the province are < 0.0001 and for job group it is < 3.63e-11

3.2 *The univariate and multivariate Cox PH model analysis*

The univariate Cox PH model analysis results indicated that all variables were statistically significant except type of exit where only one level is insignificant. Age (P < 2e-16), Hazard ratio (HR) = - 0.0361 indicating that older teachers decrease the risk of exit by 99.96. For Gender (P < 2e-16), Salary (P < 2e-16), and for Job group level B (P < 0.00607), Level C (P < 0.70395), level D (P < 1.41e-11). Three levels of the exits were highly significant (P <2e-16) and finally the province factor is also highly significant with all six levels being significant except Nyanza province.

We then used the step function in R for backward selection of variables and the Akaike information criterion (AIC). Results indicate all variables are included in the final model as shown in Table 4. These variables were then used in the multivariate analysis with the Cox PH model. According to the model results (Table 5), holding other factors constant age is highly significant and older teachers have a HR= 0.0313 and P < 2e-16), Salary (P < 2e-16) and HR = 1.000, Job group level B (P < 0.00607), Level C (P < 0.70395), level D (P < 1.41e-11). Three levels of the exits were highly significant at (P <2e-16) and finally the province factor is highly significant and all six levels are significant except Eastern province with P < 0.5494. North Eastern region has HR =1.347; Nyanza HR= 0.931 and Eastern HR = 0.987.

4.0 Discussion

4.1 Non-parametric tests

From the analysis in Table 1, we can predict 50% of the female teachers stay in the profession for 38.2 years before exit while 50 % of the male teachers stay for 36.4 years before exit. The Kaplan-Meier curve (Figure 1) also clearly shows that female teachers have a slightly higher survival rate as compared to the male teachers.

The analysis in table 2 indicated that 50% of the teachers aged 35 years and below stay for about 12 years before exit and those above 35 years old stay for about 37 years before exit. The Kaplan-Meier curve (figure 2) for younger teachers is much steeper than that for older teachers indicating that, older teachers have a higher survival rate as compared to the younger teachers.
This could be attributed to the fact that younger teachers may be more aggressive in searching for greener pastures given that they have fewer responsibilities and therefore can easily risk quitting.

The different groups of teachers classified by gender, age, job group, and province have different survival experiences as shown in our results (Table 3). The P value is < 3.63 e-11 for job group and < 0.0001 for the other covariates indicating significant differences. The log-rank test statistic (equation 2) was used to test the null hypothesis that there is no difference in the survival experience of the groups. Smaller P values provide evidence against the null hypothesis.

4.2 Multivariate Cox PH model analysis

We found all variables to be statistically significant (Table 4) and we can conclude that age, gender, job group, province, salary and type of exit are significantly associated with time to exit of teachers from service (Singer and Willet, 1988, 1991; Murnane, 1984; Ingersoll, 2001).
From the results of the model, holding other factors constant, older teachers reduce the hazard or risk to exit by 96.9% relative to the younger teachers. Young, less experienced teachers tend to have higher attrition rates compared to older more experienced teachers. This confirms earlier findings (Ingersoll, 2001; Adams, 1996).

In our study gender is significantly associated with time to exit from service (P value < 2e-16). A male teacher increases the risk to termination of service by 23 % (HR=1.228) compared to a female teacher (female teacher was used as reference). This could be attributed to the fact that men are providers for the family especially in the African social setting and therefore are looking for better opportunities.

Salary has a high significant P value (< 2e-16) indicating it is strongly associated with the risk to termination of teachers’ services. An increase in one unit (KSh) of salary decreases the risk to termination at any given time by1 fold (HR=1.00).The findings agree with previous studies (Guarino et al, 2004; Murnane and Olsen 1989 and 1990) Compensation of teachers is an important factor in retention of teachers but other factors may respond differently to salary increments (Imazeki, 2005).

Teachers enter into different job groups at the beginning of their career depending on their academic achievements and later on get promoted to upper job groups. Higher job groups (D with HR = 5.998)) tend to increase the hazard to termination by 5.998 fold, holding other factors constant. Most of the teachers in these job groups (P, Q and R) are in the age bracket of retirement while others may transfer service to non-teaching assignments in other offices.

The province of working station of a teacher is significant; Working in Eastern province reduces the risk to termination by 0.023%. However, Eastern province is not significant; while working in North Eastern parts of the country increases the risk to exit by about 35% relative to those working in Central province.

From figure1 and 2 there is no evidence of the KM curve crossing hence the proportional assumption of the model has been met. From the results of Table 4 the confidence intervals of the hazard ratios are not wide suggesting that there is no problem of over fitting. The study used a very large data hence the results are valid.

5.0 Conclusion

In this study we have presented the semi-parametric Cox PH model (equation 3) to estimate the time to exit of teachers in public schools and institutions in Kenya. This model is appropriate for this study since it can handle data with censored events. In our study the censored events are teachers who were still working at the time of data collection. The Cox PH model is widely used in the analysis of survival data since it makes no particular probability distribution assumption for the survival times, while providing reasonably good estimates of regression coefficients (Collet, 2003).The model is however based on the proportional assumption making it a semi-parametric model.

By using the Cox PH model we are able to incorporate various factors including demographic characteristics to model the time to exit of teachers from service. We are also able to determine when the event is most likely to occur.

In this study the explanatory variables age, gender, Job group, salary, Province and type of exit are significantly associated with time to exit. There were survival differences in groups of the covariates according to the log rank test. Covariate job groups had the most significant value ($P < 3.63e-11$).
Both univariate and multivariate analysis results of the Cox PH model support this. The goal of this study was to estimate the time to exit for teachers and the goal has been achieved.
We have shown how these factors affect the survival time of teachers in service. Some of these factors are responsive to policy changes. Such findings can be beneficial to the government, educational administrators and planners in coming up with better retention programs and strategies. Teachers working in hardship areas like North Eastern (HR= 1.347) increases the risk to termination of services. The government needs to give some incentives and provide proper security to teachers working in such hardship areas to encourage them to continue working in these areas.

This study has only dealt with six explanatory variables and this may not be fully exhaustive due to data limitation on teacher attrition. Data limitation on teacher attrition is more pronounced in Sub-Saharan Africa (International Task Force on Teachers for EFA 2010). One data source on teacher attrition is from the school returns through an annual written survey, which usually contain incomplete coverage. Most schools are poorly equipped and cannot come up with adequate information such as reason for exit of a teacher, destination of the departing teacher, type of exit, gender, level of qualifications and age of such teachers.
Data collected from the human resource records like the TSC provides another data source on teacher attrition; however such records do not always contain all the details of teachers like exact location of posting, subject specialization, whether they were teaching in primary or secondary schools, working conditions (International Task Force on Teachers for EFA 2010).
Guarino et al (2004) also noted the importance of reliable data on teacher retention and attrition, and importance of longitudinal data tracking of teachers’ employment.
With proper data collection methods further study on this area could be done incorporating more explanatory variables.

5.0 Acknowledgement

We are grateful to the Teachers' Service Commission for permitting us to use their data and the TSC staff at the IPPD and pension section for availing the data for this research.

References

Adams, G.J. (1996). Using a Cox Regression Model to Examine Voluntary Teacher Turnover. *Journal of Experimental Education*, 64, 267-285.

Bowman, G.D., & Dowling, N.M. (2008). Teacher attrition and retention: A meta-analytic and narrative review of the research. *Review of Educational Research*, **78**, 367 – 411.

Breslow, N.E. (1974). Covariance Analysis of Censored Survival Data. *Biometrics*, **30**, 89-99.

Charters, W.W.jr., (1970). Some factors affecting survival in school districts. *American Educational Research Journal*, Vol.7, No.1, pp.1-27.

Collet, D. (2003).*Modelling Survival data in medical research*, London: Chapman and Hall.

Cox, D.R. (1972). Regression Models and Life-Tables (with discussion). *Journal of the Royal statistical society –series* ***B***, *(*34):187-220.

Crandell, R.M., & Howell, S.J. (2009). An Analysis of teacher Choices: Transfers, Attrition and Retirement. California State University. [Online] Available:
http://www.csus.edu/indiv/h/howellj/papers/TeacherAttrition_Jan2009.pdf (Retrieved; October 24, 2013).

Education International (2007). *Teacher Supply, Recruitment and Retention in Six Anglophone Sub-Saharan African Countries*: (Researcher Dennis Sinyolo, El Coordinator, Education and Employment).

GCE/Action Aid (2005a), Contradicting Commitments, How the Achievement of Education for All is being undermined by the International Monetary Fund. London: Action Aid.

Government of Kenya (2007). Kenya Vision 2030, Ministry of State for Planning , National Development & Vision 2030 and Office of the Deputy Prime Minister and Minister of Finance, Nairobi, Kenya.

Guarino, C. M., Santibanez, L., and Daley, G.A., & Brewer.D. (2004).A review of the research literature on teacher recruitment and retention (TR-164- EDU). Santa Monica, CA: RAND.

Imazeki, J. (2005): Teacher salaries and Teacher attrition. *Economics of education Review*, 24 (2005) 431- 449.

Ingersoll, R.M. (2001). "Teacher Turnover and Teacher Shortages: An organizational analysis".*American Educational Research Journal*, fall 2001, Vol. 38, (3). 499- 534.

International Task Force on Teacher for EFA (2010).Teacher attrition in Sub-Saharan African: *The neglected dimension of the teacher supply challenge. A review of literature*, February 2010.

Kadzamira.E. C., (2006). *Teacher Motivation and incentives in Malawi*, Research Fellow Centre for Educational research and Training, University of Malawi.

Kalbfleisch, D.J., & Prentice, L.R. (2002).The statistical analysis of failure Time data, 2nd Ed. New York: Wiley.

Kaplan, E.L., & Meier, P. (1958).Non-parametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53:457- 481.

Kenya National Bureau of Statistics & Ministry of Planning and National Development. (2012): *Economic Survey* 2012 Highlights.

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

Murnane, R. J., & Olsen, R. J. (1990). The effect of salaries and opportunity costs on length of stay in teaching: Evidence from North Carolina. *Journal of Human Resources*, 25, 106– 124.

Murnane, R. J., Singer, J. D., & Willett, J. B. (1989). The influences of salaries and opportunity costs‖ on teachers' career choices: Evidence from North Carolina. *Harvard Educational Review*, 59, 325–346.

Murnane, R.J. (1984). Selection and Survival in the teacher labour market. *The Review of Economics and Statistics*, 66(3), 513- 518.

Murnane, R.J. (1981). Teacher Mobility revised. *The Journal of Human Resources*, 16(1), 3-19.

R Development Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Singer, J. D. & Willet, J. B. (1988) "Detecting Involuntary Layoffs in Teacher Survival Data: The Year of Leaving Dangerously." *Educational Evaluation and Policy Analysis*, **10** (3), 212-224.

The Ministry of Education, Science and Technology, Government of Kenya (2005). *Challenges of implementing free primary education in Kenya: Assessment repor*t, UNESCO, Nairobi.

The Ministry of Education, Science and Technology, Republic of Kenya (2005): Kenya Sessional Paper, No. 1 of 2005 on a Policy Framework for Education, Training and Research.

UNESCO - EFA (United Nations Educational, Scientific and Cultural Organization-EFA Global Monitoring Report Team). (2013/4): *Teaching and Learning: Achieving Quality For All*. Paris: UNESCO.

UNESCO - EFA (United Nations Educational, Scientific and Cultural Organization-EFA Global Monitoring Report Team) (2005a), Education for All Global Monitoring Report – *The Quality Imperative*. Paris: UNESCO.

UNESCO (2006a), Education for All Global Monitoring Report – *Literacy for Life*. Paris: UNESCO. 73.

VSO (2002): *What Makes Teachers Tick? A Policy Report on Teacher Motivation in Developing Countries*. VSO, UK.

Willet, J.B., & Singer, J.D. (1991). From Whether to When: New methods of Studying Student Dropout and Teacher Attrition. *Review of Educational Research*, **61** (4): 407-450.

APPENDIX

Tables and figures of survival time of teachers classified by gender and age.

Table 1: Kaplan-Meier summary for survival time of teachers classified by gender

Gender	Records	Events	Median (95 % C.I)
Female	139642	9806	38.2(38 – 38.4)
Male	166824	18597	36.4(36 -36.4)

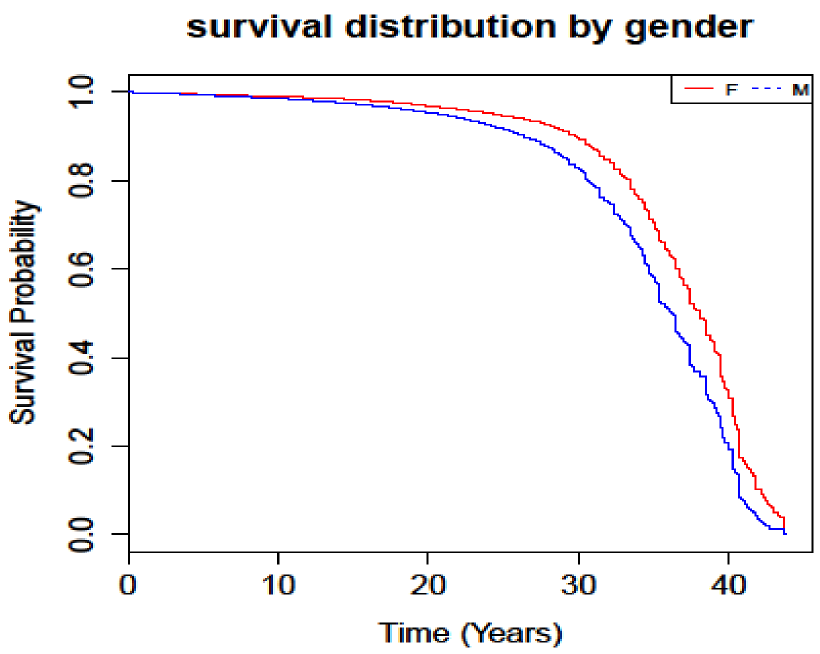

Figure 1: Survival distribution of teachers classified by gender

Table 2: Kaplan-Meier summary for teachers classified by Age (Young teachers are 35 years and below; older teachers above 35 years)

Factor	Events	Median (95 % C.I)
Younger teachers	57,471	12.3 (12.1- 12.6)
Older teachers	248,995	37.0 (37.0 – 37.4)

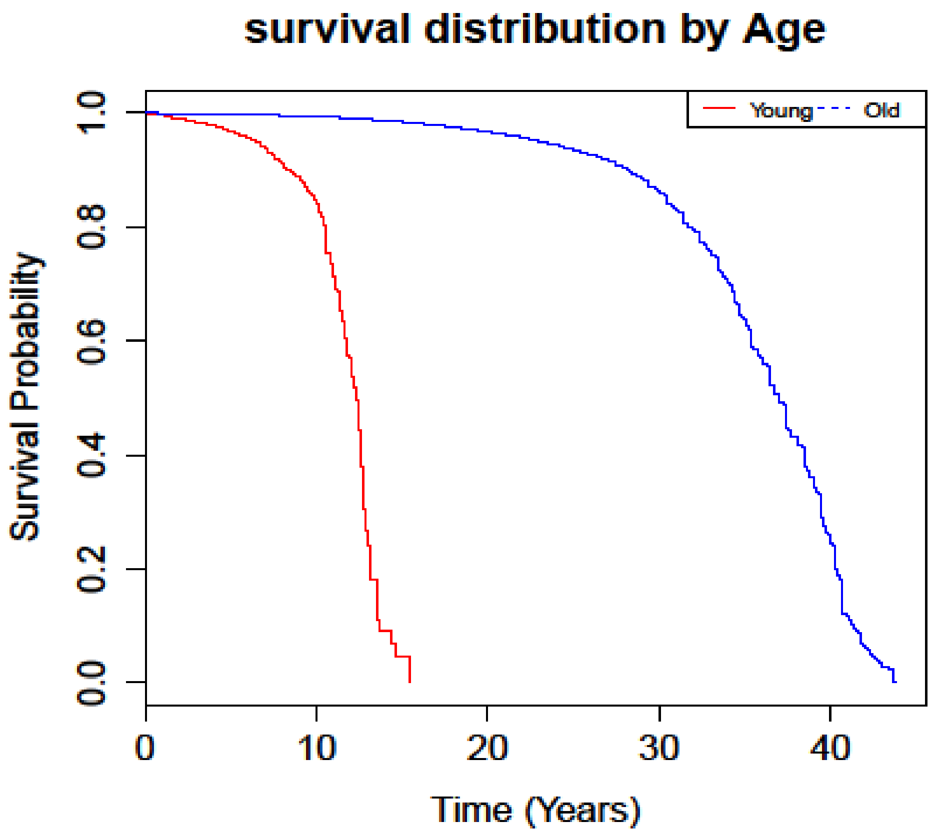

Figure 2: Survival distribution of Teachers classified by Age

Table 3: Comparison of survival curves of teachers in different groups using Log rank test

Factor	df	Chisq	P- value
Gender	1	1380	P < 0.0001
Age	1	11784	P < 0.0001
Job group	3	51.6	P < 3.63e-11
wprovince	7	5645	P < 0.0001

Table 4: model selection results using the AIC values

Step	DF	AIC
- none		508329
- wprovince	7	508421
- gender	1	508682
- salary	1	509492
- job group	3	510 392
- age	1	515 367
-exit type	4	587679

Table 5: ANOVA table of the Multivariate Cox PH model analysis

Factor	Coef (β)	HR (95%C.I)	P- value
Age			
Age (35 yrs and below)	-	As reference	
age (above 35 yrs)	-3.465	0.0313 (0.029,0.0360)	<2e-16
Gender			
female	-	As reference	
male	0.2058	1.228 (1.20 ,1.26)	<2e-16
Job group			
A	-	As reference	
B	0.8810	2.413 (2.32 – 2.51)	<2e-16
C	1.188	3.2820 (3.09 – 3.48)	<2e-16
D	1.791	5.998 (5.21 – 6.91)	<2e-16
salary	-3.277e-05	1.000 (1.000- 1.000)	<2e-16
Province stationed			
Central	-	As reference	
Coast	-0.05793	0.9437(0.891-9.99)	0.0497
Eastern	-0.01237	0.9877(0.949- 1.03)	0.5494
Nairobi	0.08452	1.088(1.03-1.15)	0.0049
North Eastern	0.2982	1.347(1.18-1.54)	1.74e-05
Nyanza	-0.07125	0.9312(0.892-0.972)	0.0012
Rift Valley	0.03999	1.041(1.00-1.08)	0.0458
Western	0.09721	1.102(1.06-1.15)	2.22e-06
Type of exit			
exit 0 (death)	-	As reference	
Exit 2 (retire mand+ contract expiry)	-1.218	0.2959(0.287-0.305)	<2e-16
exit 4 (dismissal, retire P.I)	0.5322	1.703(1.60-1.81)	<2e-16
Exit 5 (Leave service early like resignation)	-0.3562	0.7003(0.678-0.724)	<2e-16