

UNIVERSITY OF GHANA
COLLEGE OF BASIC AND APPLIED SCIENCES
SCHOOL OF PHYSICAL AND MATHEMATICAL SCIENCES
DEPARTMENT OF STATISTICS

**FACTORS THAT INFLUENCE PROMOTION OF UNIVERSITY OF
GHANA LECTURERS: A SURVIVAL ANALYSIS APPROACH**

BY
YEBOAH ERIC KWAME
(10508540)

**THIS THESIS IS SUBMITTED TO THE UNIVERSITY OF GHANA,
LEGON IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR
THE AWARD OF MPhil STATISTICS DEGREE**

JULY 2017

DECLARATION

Candidate's Declaration

I hereby certify that, this submission is the result of my own research work towards the award of an MPhil degree in Statistics and no part of this thesis has been presented for another degree in this University or elsewhere.

SIGNATURE: _____

DATE: _____

YEBOAH ERIC KWAME (10508540)

(Candidate)

Supervisors' Declaration

We hereby certify that this thesis was prepared from the candidate's own work and supervised in accordance with guidelines on supervision of thesis laid down by the University of Ghana.

SIGNATURE: _____

DATE: _____

DR. KWABENA DOKU-AMPONSAH

(Principal Supervisor)

SIGNATURE: _____

DATE: _____

DR. ISAAC BAIDOO

(Co-Supervisor)

ABSTRACT

Survival analysis has been invaluable in studies involving time to an event. The methods encompassing the idea of Survival Analysis appears to be commended by most researchers particularly in the medicinal, engineering, agricultural and actuarial fields.

This paper applied survival methods to promotion data of lecturers in the University of Ghana. The main objective was to identify the factors prominent in facilitating an early promotion of University of Ghana lecturers. The proportionality assumptions were assessed and some model diagnostics were made with some graphical presentations. The proportionality assumption was satisfied and the models were compared by their AIC values. The Weibull Proportional Hazard model recorded the lowest AIC value making it a better model to fit the data set.

The average survival time for promotion is 78.847 months with a standard deviation of 3.72 and a 95% CI [71.497, 86.198]. This means that it take an average of 6.57 years for a lecturer to earn their first promotion.

This study considers the variables *age, number of UG committee membership, number of international conferences attended, number of technical reports written, number of working papers, number of books published, qualification before entry and origin of Master's certificate* as significant at 5% significance level, in influencing the time until a lecturer gets promoted at the University of Ghana.

The variables, *number of children, marital status, gender, national committee membership, and college of affiliation* were not significant at 5% significance level.

Evidently, young lecturers are more likely to be promoted as compared to aged lecturers. But the higher the number of UG committee membership, International conferences attended, books published and technical reports written, the more likely it is to get a promotion.

DEDICATION

I dedicate this thesis to my foster and God-parents, Mr. Charles Opong Otchere and Mrs. Stella Otchere, their children, Gifty and Bright; my siblings, Blessing and Matilda and my beloved parents, Mr. Peter Kwasi Yeboah and Madam Margaret Afia Nyarkoaa.

ACKNOWLEDGEMENT

I thank the Almighty God for the privilege, knowledge and the resources to pursue my education up to this level. I will also like to show appreciation to my principal supervisor and co-supervisor Dr. Kwabena Doku-Amponsah and Dr. Isaac Baidoo respectively, without them this work would not have been a worthwhile. I sincerely thank them for the guidance, patience, dedication and the detailed scrutiny of this thesis. I render my heart-felt gratitude to the Head of Statistics Department, University of Ghana, Dr. K. Doku-Amponsah for putting together the requisite structures to help research work. My appreciation to Dr. Samuel Iddi who sometimes stepped in to guide me on my thesis trajectory anytime my assigned supervisors were unavailable.

Much gratitude also goes to a good friend of mine, Crystal Bubune Letsa, for her inspiration and motivation. I couldn't have had the impetus to complete this work without her support.

My course mates, you were all awesome. I appreciate each and every one of you.

To my dearest Mabel, my long-standing friends Sampson Baidoo, Daniels Dorsey, Agnes and my entire family, I say a big thank you to all for your care, inspiration, assistance, patience and prayers throughout my study period and may the God Almighty bless you all immeasurably.

Thank you all and God bless you exceedingly.

TABLE OF CONTENTS

DECLARATION.....	i
ABSTRACT.....	ii
DEDICATION.....	iv
ACKNOWLEDGEMENT.....	v
CHAPTER ONE.....	1
INTRODUCTION	1
Introduction	1
1.1 A Lecturer and Promotion at the University of Ghana	1
1.2 Research Problem Statement.....	2
1.3 Research Questions	3
1.4 Statistical Methodology.....	3
1.5 Objectives of Study	4
1.6 Significance of This Study	4
1.7 Scope of Study	5
1.8 Organisation of Work.....	5
CHAPTER TWO	7
LITERATURE REVIEW	7
2.1 Job Mobility	7
2.2 Promotion of Workers.....	8

2.3	Promotion of Lecturers in the University of Ghana.....	9
2.4	Applications of Survival Analysis.....	9
2.5	Survival Analysis and Promotion of Workers.....	9
CHAPTER THREE		11
METHODOLOGY		11
3.1	Study Area and Experimental Design	11
3.2	Ethics.....	12
3.3	Statistical Techniques.....	13
3.3.1	Overview of Survival Analysis.....	13
3.3.2	Censoring and Truncation	15
3.3.3	Analysing Survival Data.....	17
3.3.4	Proportional Hazard (PH) Modeling	27
3.3.5	AFT Modeling.....	33
3.3.6	Diagnostic Tests	36
3.3.7	Modeling and Model Selection.....	39
3.4	Statistical Software.....	39
3.5	Dissemination of Results.....	39
CHAPTER FOUR.....		40
RESULTS AND DISCUSSION		40
Introduction.....		40

4.1 Explanatory Variables	40
4.2 Preliminary Analysis	41
4.3 Modeling Process and Analysis	47
CHAPTER FIVE	55
CONCLUSION AND RECOMMENDATIONS	55
Introduction	55
5.1 Conclusions	55
5.2 Recommendation.....	57
5.3 Future Research.....	57
REFERENCES	58
APPENDIX A.....	62
APPENDIX B.....	72
APPENDIX C.....	74

LIST OF TABLES

Table 4.1 Variable Coding for Categorical Variables.....	41
Table 4.2 Variable Description for Categorical Variable	43
Table 4.3 - Continuous variable descriptive.	44
Table 4.4 Log Rank and Wilcoxon (Breslow) test for equality of survivor functions	46
Table 4.5 Proportional hazard assumption test for the covariates	49
Table 4.6 Comparing PH models using Log-likelihood and AIC	50
Table 4.7 Comparing AFT models using Log-likelihood and AIC	50
Table 4.8 Standard errors of Exponential, Weibull, Log-logistic and Lognormal AFT models.....	51
Table 4.9: Weibull Proportional Hazard Model	54
Table C1: Exponential Proportional Hazard Model	66
Table C2: Gompertz Proportional Hazard Model.....	67
Table C3: Exponential Accelerated Failure Time Model	68
Table C4: Weibull Accelerated Failure Time Model	69
Table C5: Lognormal Accelerated Failure Time Model.....	70
Table C6: Log-logistic Accelerated Failure Time Model.....	71

LIST OF FIGURES

Figure 4.1	Kaplan-Meier Curve	62
Figure 4.2	Kaplan-Meier Curve comparing Gender	62
Figure 4.3	Kaplan-Meier Curve comparing Colleges	63
Figure 4.4	Kaplan-Meier Curve comparing Marital Status	63
Figure 4.5	Kaplan-Meier Curve comparing Qualification Before Entry (Masters and Phd)	64
Figure 4.6	Kaplan-Meier Curve comparing Local or Foreign Masters degree.....	64
Figure 4.7	Component Bar Chart Comparing Survival Frequencies	65

CHAPTER ONE

INTRODUCTION

Introduction

This study uses Survival Analysis to decide which variables influence the time a lecturer in the University of Ghana is promoted. The organisation of this chapter is as follows: it begins with a short background of the research topic, and afterward continues to investigate the general profile of the study area, the problem statement, research questions and targets. Research methodology, as well as scope and constraints of the study are additionally talked about.

1.1 A Lecturer and Promotion at the University of Ghana

A lecturer at the University of Ghana is a senior member who looks into, educates and might be in charge of some managerial and administrative roles. Lecturing as an occupation can be an exceedingly remunerating profession, with the opportunity to research in one's picked area of intrigue, and the likelihood of having a high effect in scientific applications for instance.

Lecturers may explore as people or work in a group, which might be multidisciplinary as well as include coordinated effort crosswise over colleges and even nations. Here and there they work exclusively on issues that intrigue them specifically. Regularly, they supervise research students or postdoctoral researchers.

The major addition to the purpose of a scholarly employment is the relative flexibility to set your own particular working hours, and your research motivation. The employment can be exceptionally fulfilling, with the chance to impact the bearing of research in your field. It can likewise be changed, with time split between an assortment of educating, investigating and extracurricular exercises.

A down point will be workload that can regularly be high. In more senior positions, the time accessible for research might be decreased because of regulatory and supervisory obligations. Marking exams and assignments can be monotonous particularly for bigger classes and a few courses particularly those including more practical sessions. Additionally, the pressure of churning out research output in order to attract funding can be daunting; prominently referred to among the University people group as to either 'publish or perish'.

Presentations about researches at seminars and conferences, and write ups in recognized journal articles are well-known factors which enable lecturers to be promoted. Being promoted is an elevation in status or position. However there might be other contributory components which may impact a lecturer's promotion, which this study seeks to unravel.

1.2 Research Problem Statement

Promotions are of significance to each employee since it is an indication of affirmation for diligent work and gives the worker some level of esteem, benefits and now and again, included obligations. Establishments may either have one of a kind statutes or be utilizing another organisation's statutes as reference. The promotion of lectures in every institution is reliant on the overseeing statutes. There are however a few necessities regular to every one of the organisations. There may be fixed time to promotion or variable timing. There are contributory factors to a person's promotion whether the time is fixed or variable.

Admittedly, promotions in the University of Ghana is an issue on the mind of already employed lecturers, a civil argument among yearning lecturers and a reference to other higher institutions of learning.

Dominantly, this research seeks to uncover the issue of how long it will take a lecturer to gain his/her first promotion and investigate the components or factors that influence or feed into earning a promotion.

1.3 Research Questions

Questions that would be answered in this research are:

1. At what time, from the time of entry into lectureship, does a University of Ghana lecturer earn their first promotion?
2. What factors predispose a University of Ghana lecturer to a faster promotion?

1.4 Statistical Methodology

Survival analysis is a set of statistical methods for analysing the timing and occurrence of an event (Allison, 1995). Survival Analysis concerns itself with time to occasion information. In the broader sense, it comprises of systems for non-negative valued random variables, for example, time to death, time to relapse of an illness, length of stay in a health centre, duration of a strike action, money paid by health insurance, viral load measurements, time to finishing a dissertation or thesis and so on. “Territories of survival studies include: clinical trials, prospective cohort studies, retrospective cohort studies, retrospective correlative studies.” (Ibrahim, 2014).

1.5 Objectives of Study

- ✓ Determine the survival rate of lecturers' promotion with the application of non-parametric survival data analysis.
- ✓ Compare parametric and semi-parametric PH models by their AIC values to select the best fitted model to the data.
- ✓ Highlight factors that contribute to the early promotion of University of Ghana lecturers.
- ✓ Make recommendations to staff and management regarding promotions of University of Ghana lecturers.

1.6 Significance of This Study

This study is of great significance because:

- It will give up and coming lecturers and understanding on the best way to adjust, set needs, their obligations and where their concentration should lay keeping in mind the end goal of climbing the ladder in the scholarly community.
- It will fill in as a form of reference point for other higher institutions of learning on where their objectives ought to be, particularly when recruiting lecturers.
- It will help the university to consider her approach on promotions if it emerges from this study that the current one is not effective.

1.7 Scope of Study

The study focuses on determining how long it will take a lecturer of the University of Ghana to gain his/her first promotion and investigate the components or factors that influence or feed into earning a promotion.

- Proportional stratified computations were used to determine the number of lecturers to be involved from each college.
- Convenient random sampling method was then employed to collect primary data from a hundred and fifty-one (151) full-time lecturers from the University of Ghana across all four colleges, having administered 200 questionnaires, a copy of which can be found in Appendix C. It became necessary to use convenient random sampling because it became increasingly challenging to get in touch with lecturers who had been originally listed as respondents.

The fundamental procedure is the Survival Analysis since it accounts for censored data. Different techniques of Survival Analysis are explored, specifically, Proportional Hazard (PH) modeling and the Accelerated Failure Time (AFT) modeling.

The PH models include: Cox Regression, Weibull, Gompertz and Exponential models while the AFT models are: Exponential, Weibull, Log-logistic and Lognormal models. STATA was the statistical software used for the analysis.

1.8 Organisation of Work

The research work has been organized in this thesis as follows:

Chapter one gives an introduction to the study.

Chapter two reviews the related literature on the subject.

Chapter three describes the theory of the model to be used, plans and strategies for obtaining solution.

In *Chapter four* we deal with the data collection method and the data is analysed and displayed.

This section likewise contains the translation and discussion of the results.

Finally, in *Chapter 5* we conclude the whole study and state particular suggestions to stakeholders in light of the real discoveries made in the study

CHAPTER TWO

LITERATURE REVIEW

Literature that studies University of Ghana lecturers were limited in scope. However, studies on job promotion and the influencing factors thereof were expansive. The study of occupational mobility cuts across all fields of human endeavour. Literature therefore covers, expansively, the methods of productivity assessment, the factors of promotion and determining the factor of significant influence by using a relatively superior method of Survival Analysis.

2.1 Job Mobility

Job mobility is the idea used to portray any movement on a job. For employees, job mobility is gives some satisfaction on an individual level. It affords them a chance to grow new abilities, expand their understanding and develop proficient systems. As such findings from studies on time to promotions and the factors that feed into that are quintessential in every employer or employee's occupational itinerary. The following sorts of job mobility can be recognised:

- i. Occupational/vocation mobility: when a worker completes an alternate job (occupational category) for a similar business (e.g. moving from junior to senior supervisor).
- ii. Intra-sectorial mobility: when a worker does a similar job for another business in a similar segment (e.g. moving as a post-doc scientist starting with one college then onto the next).
- iii. Inter-sectorial mobility: when an employee does a similar job for another employer in another sector (e.g. moving from college to industry).

Internal labour markets are clusters of jobs that have three fundamental auxiliary elements: job ladders, entry at the bottom, and internal promotion based on knowledge or skill development.

This research is solely going to be on transition of a worker from one level to another defined above as occupational or career mobility (Althauser, R. & Kalleberg, 1981).

2.2 Promotion of Workers

The determinants of mobility may rely upon the definition received; promotions are separated by whether or not they are related with an adjustment in the assignments performed. There are existing financial speculations which have tended to systems creating work mobility inside and between firms. One model assumes that in the labour market, employers determine wages and employment (vacancies) and predicts that workers change jobs in response to differences in wage rates. In the event that specialists were the same, at that point firms will effortlessly think that it's simple to replace a worker. While profitability is expanded hardly by human capital, a few firms require particular abilities keeping in mind the goal is efficiency and productivity. Consequently, worker turnover turns into a matter of significance since a few firms prepare a few representatives with the intent to get them the expertise for the company's advancement. The organisations, generally private ones, may build wages of exceptionally talented workers with a specific objective to hold them. In order to increase profitability of workers, employers may also create a promotion scheme which will have an associated wage and privileges to motivate workers in firm-specific human capital (Ferreira, 2009).

Promotions can be viewed in the sense of competitions. A promotion is a prize that is allocated to workers who out ranks every single other worker in an organisation over a given period. The prospects of promotion in itself is incentivising as it induces one to exert effort, and “winners” are moved to higher positions that involve for instance, an elevated status, higher responsibility, or higher remuneration. (Bognamo, 2001; Lazear, 1981).

2.3 Promotion of Lecturers in the University of Ghana

The University of Ghana statutes gives the strategy for the promotion of lecturers. Also, regardless of the level one is qualified to apply for; an aspiring applicant must download and fill an online application form, found on the University's website. This form gives information on what elements are examined for promotions. The application goes through a series of processes. *Schedule F* of the **Basic Laws of The University of Ghana, 2012** has some additional information for evaluation and promotion in the University of Ghana.

2.4 Applications of Survival Analysis

Every one of the cases above has information comprising censored observations in which the end occasion has not occurred in each observation or when information on a case is only known for a limited duration. The censoring time is the main information to find cumulative survival probability in the survival analysis. In other words, the great advantage of using survival analysis is to analyse censored cases in analysis. Another quality of utilizing survival analysis is the utilization of covariates and the capacity to assess the magnitude of specific influences which can be analysed (e.g., age, level of instruction) (Wang, Little and DelHomme-Little, 2012).

2.5 Survival Analysis and Promotion of Workers

Survival analysis is a tool that has been used in the demographic and organisational literature to examine concepts such as population ecology, absenteeism and turnover. (Lee, et al., 2008; Morita, et al., 1989, 1993)

This approach is preferable in examining the effect of different types of boundary crossings on the likelihood of career progression because it is designed to model “time to event” data and is well suited for data at different time points with dependent variables (Morita et al., 1989).

A comparison with the logistic and OLS regression gives an indication that survival analysis is a befitting tool for us to use for this study because the probabilities of the event of an occasion (career promotion) are depicted as an function of time, and it enables the probabilities to vary starting with one time point then onto the next (Morita et al., 1993).

The response variable is the time to promotion in career. The predictor variables included sex (0 = *female*, 1 = *male*), marital status (0 = *not married*, 1 = *married*), and level of education as control variables because earlier studies have given an indication sex and marital status may be influential in career promotion (Foster N., 2001; Fernandez, J. P., 1981).

A study of data combined from multiple sources found that sex is a key socio-demographic variable that predicts career success, stating that women’s career progression is relatively slower to that of men. Marital status has been appeared to impact vocation progression, with married people accomplishing more positive profession results than unmarried people (Judge & Bretz, 1994; Ng et al., 2005; Pfeffer & Ross, 1982). One’s level of education, as a type of human capital phenomenon, has also been found to have statistical significance with profession results. (Ng *et al.*, 2005; Pfeffer & Ross, 1982).

CHAPTER THREE

METHODOLOGY

In this chapter, we present the statistical methodology for analysing the promotion of University of Ghana lecturers. The principal statistical procedure used in this work is survival analysis. In addition, the study area and design of the experiment are presented.

Time to events encompass survival time in non-promotional and promotional events. The survival time until promoted in an organisation is an example of time in a promotional event. Conversely, the period after graduation until one gets a job can be categorized under survival time in non-promotional event. An example of the application of survival analysis for non-promotional event is ascertaining the determinants of the survival of a musical icon on a well-branded record label, where time is an important variable for the development of the label.

3.1 Study Area and Experimental Design

The study includes presently employed full time lecturers in the University of Ghana. The table below shows the experimental design of the study

Data points	Description
1. Date of employment	The date a lecturer is employed
2. Level of employment	The level a lecturer was employed whether as an assistant lecturer, lecturer, senior lecturer, etc.
3. Qualification	The qualification a lecturer had as at time of employment.
4. Nationality	Country of birth
5. Region of origin	
6. Gender	
7. Religion	
8. Age	The lecturer's date of birth or age as at time of employment.
9. Schools attended	Names of tertiary institutions a lecturer attended and countries where the schools are located.
10. Promotions and dates	The date a lecturer gets promoted.
11. Publications	The number of publications a lecturer submits in order to earn a promotion.
12. Marital status	The marital status is preferred with two levels single or married with date of marriage specified.
13. Children	Number of children and dates of births if available. The names of children are to be excluded.

3.2 Ethics

There is the requirement for high moral benchmarks in any study. The privacy of the lecturers whose information were used were protected throughout the study.

3.3 Statistical Techniques

In this section, we present the concepts, formulae and models that were employed in analysing the “time to event” data obtained from the respondents for the purposes of determining the time until promotion and the factors that influence same.

3.3.1 Overview of Survival Analysis

Survival analysis is a study of time-to-occurrence of an event of interest. The study of time-to-event deals with data with exceptional elements. Some of these relevant concepts inherent in the study are censoring and truncation, and these require special attention when dealing with such data.

There are four important functions in survival analysis, namely:

- Survival function
- Hazard function
- Cumulative hazard function
- Mean residual life
- Density function

Survival function

Survival function ($S(t)$) is the probability of an individual surviving at least time, t . It is written mathematically as;

$$S(t) = P(T > t) \quad (1)$$

$$= \int_t^{\infty} f(t)dt \quad (2)$$

where,

$f(t) = -\frac{dS(t)}{dt}$ is the density function.

The Density Function

Another function worthy of note in survival analysis is the probability density function; for a continuous random variable and probability mass function; for a discrete random variable.

The continuous distribution T has a cumulative distribution function given as

$$F(t) = P[T \leq t], t \geq 0$$

which yields the density function when differentiated, giving

$$f(t) = \frac{dF(t)}{dt} \Rightarrow F(t) = \int_0^t f(u) du \quad (3)$$

The density function is related to the survival function by $S(t) = P[T > t] = 1 - F(t)$, where T is the survival time; $S(t)$ is the probability that a randomly selected person will survive to at least time t (the survival function)

The survival function is a monotonically decreasing function; no matter the distribution, as time increases, the chance of survival decreases, i.e.

$$S(0) = 1 \text{ and } \lim_{t \rightarrow \infty} S(t) = S(t) = 0.$$

Hazard Function

The hazard function is the momentary rate at which an event happens given that the event has not occurred at time t . The hazard function is given as

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P[t < T < t + \Delta t | T \geq t]}{\Delta t} \quad (4)$$

The hazard function gives more information about mechanisms underlying failure than the survival function and for a continuous random variable, the hazard function is obtained by;

$$h(t) = \frac{f(t)}{S(t)} = \frac{-d\ln[S(t)]}{dt} \quad (5)$$

Cumulative Hazard Function

The cumulative hazard $H(t)$ is the total of hazard rates over a period of time say, t .

$$H(t) = \int_0^t h(u)du = -\ln[S(t)] \quad (6)$$

Hence the survival function can be deduced from the relations above.

$$S(t) = \exp[-H(t)] = \exp\left[-\int_0^t h(u)du\right] \quad (7)$$

3.3.2 Censoring and Truncation

Censoring is said to be present when information on time to outcome event is not available for all study participants. Participant is said to be censored when information on time to event is not available due to loss to follow-up or non-occurrence of outcome event before the trial end (Singh, et al., 2011). The random occurrence must be free of the event of interest. Censoring generally happens when a person does not encounter the event before the study by the researcher ends, the subject drops out as study advances (attrition) or the subject is not accessible for a subsequent test.

There are three types of censoring: left, interval and right censoring.

- Left censoring: When the event of interest has already occurred at the time of observation but the exact time is not known.
- Interval censoring: This type of censoring occurs when a random variable of interest is known only to lie within an interval instead of being observed exactly.

- Right censoring:

Right censoring occurs when a subject leaves the study before an event occurs, or the study ends before the event has occurred (Lawless, J. F., 2011).

Right censoring has three types. The fixed types 1 and 2 and random censoring.

- *Fixed type 1 censoring*: in this type of censoring, a censored subject does not experience the event of interest because the experiment is set to end after “C” years of follow-up. Hence any subject who does not experience the event before the end of the experiment and follow-up time elapses is censored.
- *Random censoring*: This type of censoring occurs when subjects have different censoring time even though the experimental design has a fixed study time.
- *Fixed type 2 censoring*: In this type of censoring, the experimental design is such that, there are a pre-specified number of events of interest. Hence any subject that does not experience the event before the required number is obtained is censored.

Truncation

Truncation on the other hand, is an incomplete observation pertaining to a subject whose non-occurrence is due to the form of selection inherent in the design of experiment employed.

The types of truncation include left and right truncation. Unlike censoring which has partial information on the subject, in truncation the researcher has no information on the subject.

- Left truncation: Individuals whose event time is greater than some truncation threshold are observed, otherwise they are truncated.
- Right truncation: The inclusion of subjects with time-to-event lesser than a truncation threshold in a study

In order to have an effective methodology, the cause of both truncation and censoring should be independent of the failure of event of interest.

If present in a dataset, truncation and censoring must be included in the analyses.

3.3.3 Analysing Survival Data

Survival data can be analysed in three main ways; the *parametric*, *semi-parametric* and *non-parametric* approaches. The Kaplan-Meier (KM) estimation is an example of a non-parametric modeling of the survival function. The Cox Proportional Hazard is an example of the semi-parametric model and the Gamma, Exponential and Lognormal distributions are some examples of parametric models.

In the parametric approach, the survival time is assumed to follow a particular parametric distribution. The data must initially be tested to ascertain assumptions of the particular distribution. These tests are powerful but may lead to inaccurate conclusions if they are used inappropriately. However, in the nonparametric modeling, assumptions are not stringent and therefore are more flexible. The semi-parametric tests are flexible which make some assumptions about time-to-event data relatively easier to be tested and also model the effects of covariates. These survival estimation techniques are however able to estimate censored or non-censored data.

The two main estimation methods employed in this work are estimation using semi-parametric and parametric models. The semi-parametric model is the Cox Proportional Hazard (PH) regression model. After the Cox PH has been tested, the best-fitting parametric model out of the three is in turn compared to the Cox PH model in order to know through which method of modeling the data produces a better fitting model. Thus, it is proper to use parametric models which use the same method of modeling as the Cox PH regression model.

Accelerated Failure Time (AFT) and Proportional Hazard (PH) methods

The AFT's are usually linked to the survival function. They measure the covariates' effects which are multiplicative on the survival times. The covariates either accelerate or decelerate the occurrence of an event. Also, AFT's are assumed to be constant for any survival percentile. Unlike the AFT models, the PH models are associated with the hazard function. It measures the risk of the occurrence of an event of interest. The determinants are multiplicative to hazards and either increase or decrease hazards.

The AFT and PH models have their unique limitations. In the PH model for instance, the model tends to be restrictive and it is not easy to interpret due to its non-linear nature. This non-linearity makes it not easy to analyse the factors on especially long-term effects. Log-logistic, Lognormal, Gamma and Inverse Gaussian distributions are parametric distributions which use the AFT method of analysing survival data.

In the sections that follow, we discuss the types of models; non-parametric, semi-parametric and parametric and the techniques of modeling; PH and AFT.

3.3.3.1 Non Parametric Model: Kaplan-Meier (KM)

Most of the descriptive tests are non-parametric tests because they are relatively easier to understand and interpret. The first step is to obtain the survival times for the data.

Afterwards, the Kaplan-Meier (KM) also known as the Product Limit estimator, a non-parametric modeling approach is used to estimates the survival function.

Let $d(t)$ denote number of events (promotions) at time t . $d(t)$ is either 1 if a lecturer has earned a promotion or 0 otherwise, but there is an allowance for the possibility of ties in which the number of promotions is greater than one.

The number of individuals at risk at time t is denoted by $n(t)$. The KM survival function is given as:

$$\hat{S}(t) = \prod_{t \leq t_i} \left(1 - \frac{d(t)}{n(t)}\right) \quad (8)$$

Right censoring is accommodated in the Kaplan Meier method which is a right continuous step function which has a jump at the “death” times. After obtaining the survival functions, a graph of the survival function against time is plotted to enhance understanding of the function.

Again, the survivorship of categorical variables can be assessed by various tests, a number of statistical tests have been proposed to answer this question, and most software packages provide results from at least two of these tests. However, comparison of the results obtained by different packages can become confusing due to small but annoying differences in terminology and methods used to calculate the tests (Hosmer, Lemeshow & May, 2008).

Some of the tests include, Wilcoxon test weights, Log Rank Tests, Tarone Ware weights, Peto-Prentice weights and other tests peculiar to certain software packages.

However, in this work, the Log Rank test is used to assess differences in survivorship across categories because of its availability in the **STATA** package and the ease in result interpretation.

The Log Rank test can also be extended to covariates with more than two categories. The null hypothesis is generally given by;

H_0 : There are no differences between survival curves.

Like the KM test, this test has same assumptions. These are;

- a. the survival probabilities are the same for subjects recruited early and late in the study,
- b. censoring is unrelated to chances of survival,

- c. the events happened at the times specified.

Test statistic for the Log rank test is obtained in similar to the chi-square test procedure. The observed survival and the expected survival cell counts over categories.

$$\text{Long rank statistic} = \frac{(O_2 - E_2)^2}{\text{var}(O_2 - E_2)}$$

for a two category covariate, where O is the observed survival count and E is the expected survival count. Hence the test statistic can be approximated as

$$\text{Long rank statistic} = \sum_{i=0}^n \frac{(O_i - E_i)^2}{E_i} \quad (9)$$

This approximation formula applies to covariates with more than two categories.

The hypothesis for this test, given for instance for the covariate gender is given as;

H_0 : There is no significant difference between the survival of males and females.

H_1 : There is a significant difference between the survival of males and females.

Tested at a 0.05 level of significance, a resulting p-value lesser than 0.05, implies that, survival is different for males and females.

The hypothesis is set in like fashion for the test of survivorship for the rest of the covariates.

3.3.3.2 Semi-parametric Regression Model: Cox Regression

Cox Regression modeling is usually used for data which have subjects in groups with additional characteristics that may affect outcomes. The variables may be used as covariates which may be explanatory variables, confounders, risk factors or independent variables. The Cox PH model is not entirely a parametric model because even though its survival time has a parametric regression structure, its dependence on time is left unspecified (Hosmer, et al., 2008).

It is a semi-parametric model as the distribution of the outcome can still be unknown even if the regression parameters (the betas) are known.

The Cox regression model is one of the most used modeling methods in survival analysis because it is robust; a close approximation of the results for the correct parametric model can be done basing on the results from using the Cox model. Also, Cox model always yields a non-negative hazard rate. The exponential part of the hazard function is appealing because it causes the fitted model to always give estimated hazards that are non-negative. Another appealing property of the Cox model is that, it is still possible to estimate the β 's in the exponential part of the model even though the baseline hazard part of the model is unspecified. The hazard ratio, which is measure of effect, is calculated without having to bother about the baseline hazard function.

The final reason for the popularity of the Cox model is that, it uses more information – the survival times – than the logistic model, which considers a (0,1) outcome and ignores survival times and censoring (Kleinbaum & Klein, 2005).

The general form of the Cox model is given by;

$$h(T, x) = h_0(t) \exp \left(\sum_{i=1}^k \beta_i X_i \right) \quad (10)$$

where $h_0(t)$ is the baseline hazard and the β_i 's are the parameters with $i = 1, 2, 3, \dots, p$

The parameters in the equation are estimated using partial likelihood. This procedure takes into consideration censored data. The likelihood function is the product of the number of failure times say k .

$$L(\beta) = \prod \frac{\exp X_i \beta}{\sum_{Y_j \geq Y_i} \exp X_j \beta} \quad (11)$$

The partial log likelihood is then given by:

$$l(\beta) = \text{Log } L(\beta) = \sum \left\{ X_i \beta - \log \left(\sum_{Y_j \geq Y_i} \exp X_j \beta \right) \right\} \quad (12)$$

After forming the likelihood function, it is maximised (partial derivatives of log L are obtained with respect to the parameters in the equation) then the resulting equation is solved in order to derive valid partial MLE's of β . The partial likelihood is effectual if no two subjects experience the event simultaneously (presence of ties). If ties occur, some permutations are done to cater for the situation and hence the uses of approximations like the Breslow or Efron's approximation to partial log-likelihood.

In Cox models, the basic theory is that the hazard rates for different values of the covariates are proportional. Regression coefficients that can be used to determine the percent change in the hazard rate (i.e., the increase or decrease in the probability of a promotion) given changes in the independent variable (e.g., Age, marital status, number of publications) are generated by the Cox model (Zheng, et al., 2011). The hazard ratio is the ratio of hazard of one individual in a study to that of another individual comparing some covariates.

It is written as,

$$HR = \frac{h(t, X^*)}{h(t, X)} = \frac{h_0(t) \exp(\sum_{i=1}^k \beta_i X_i^*)}{h_0(t) \exp(\sum_{i=1}^k \beta_i X_i)} \quad (13)$$

which becomes,

$$HR = \exp[\sum_{i=1}^k \beta_i (X_i^* - X_i)]$$

because the baseline hazards cancel out each other.

3.3.3.3 Parametric Models

Exponential

The exponential distribution is a frequently used parametric model; it has one parameter and a constant hazard. Its probability density function is given by,

$$f(x) = \lambda e^{-\lambda x}, \text{ where } \lambda, x > 0 \quad (14)$$

The survival function and the hazard function of the exponential distribution are:

$$\begin{aligned} S(t) &= \exp(-\lambda t) \\ h(t) &= \lambda \end{aligned} \quad (15)$$

The exponential distribution is mathematically tractable due to its memory-less property given by,

$$P(T \geq t + z | T \geq t) = P(T \geq z), \quad (16)$$

This property makes the use of the exponential distribution limited in application fields like in industrial and health research. The no memory property simply means that, the occurrence of a future event does not depend on an event that has already occurred.

Hence the mean residual life is constant and given by,

$$E(T - t | T > t) = E(T) = 1/\lambda$$

Weibull

The Weibull distribution is used in parametric modeling of survival data analysis and has two parameters. This distribution can reduce to an exponential distribution and it is very flexible.

The parameters are λ and α , the scale and shape parameter respectively.

The shape parameter α can take different values which in turn affects the shape of the hazard function. When $\alpha > 1$, the hazard function is increasing, $\alpha = 1$ implies the hazard is constant and similar to the hazard in an exponential model and when $\alpha < 1$, the hazard is a decreasing function. Thus the parameter α is referred to as the shape parameter.

The product of the survival function and hazard function gives the density function.

$$S(t) = \exp(-\lambda t^\alpha)$$

$$h(t) = \lambda \alpha t^{\alpha-1}$$

$$f(t) = S(t) \times h(t)$$

$$\exp(-\lambda t^\alpha) \lambda \alpha t^{\alpha-1} \text{ where } \alpha, \lambda > 0 \text{ and } t \geq 0 \quad (17)$$

There are some assumptions that must be tested to ascertain the appropriateness of Weibull distribution. Even though the assumptions in other distributions differ depending on whether it is AFT or PH, for the Weibull distribution, if the PH assumption holds then the AFT assumption holds and vice versa.

Gompertz

The Gompertz model like the Weibull is also a two-parameter model; it is a parametric model that uses PH modeling but not AFT. The Gompertz is popularly used in modeling mortality curves. The hazard function of the model is given by,

$$h_0(t) = \theta e^{\alpha t} \quad (18)$$

and survival function

$$S(t) = \exp\left[\frac{\theta}{\alpha}(1 - e^{\alpha t})\right] \quad (19)$$

thus the density function is;

$$f(t) = \theta e^{\alpha t} \exp\left[\frac{\theta}{\alpha}(1 - e^{\alpha t})\right] \text{ where } \alpha, \lambda > 0 \text{ and } t \geq 0 \quad (20)$$

Log-Logistic

The log- logistic is a parametric modeling distribution that strictly models using the AFT approach of modeling in Survival analysis. The hazard function is not monotonic, it is given by:

$$h(t) = \frac{\lambda p t^{p-1}}{1 + \lambda t^p} \quad (21)$$

where p is the shape parameter and λ the scale parameter.

For $p \leq 1$, the hazard decreases and for $p > 1$, the hazard is uni-modal.

As the name suggests, the log-logistic is a proportional odds survival model and therefore has the underlying assumption that, proportional odds (PO) have odds ratio constant over time.

The definitions are thence given in terms of odds.

Survival odds, is the odds of surviving beyond a time t .

$$\frac{S(t)}{1 - S(t)} = \frac{P(T > t)}{P(T \leq t)} \quad (22)$$

$$S(t) = \frac{1}{1 + \lambda t^p}$$

and

$$1 - S(t) = \frac{\lambda t^p}{1 + \lambda t^p}$$

Failure odds, is the odd of experiencing the event by time t ; this is given by the reciprocal of equation (22) as:

$$\frac{1 - S(t)}{S(t)} = \frac{P(T \leq t)}{P(T > t)} = \frac{\left(\frac{\lambda t^p}{1 + \lambda t^p}\right)}{\left(\frac{1}{1 + \lambda t^p}\right)} = \lambda t^p \quad (23)$$

The log-logistic has its survival function as;

$$S(t) = \frac{1}{1 + \lambda t^p} = \frac{1}{1 + (\lambda^{1/2} t)^p} \quad (24)$$

Lognormal

The Lognormal is also a parametric distribution that is AFT, not proportional odds and its survival and hazard can just be expressed in terms of integrals. Lognormal and log-logistic distributions are similar and result in similar models. The survival time t is such that, $t \log$ follows a normal distributed with mean μ and variance, 2σ . The Lognormal distribution is high in popularity partly due to the fact that the cumulative values of $y = \log t$ can be gotten from the tables of the standard normal distribution and the matching values of t are then obtained by taking antilogarithm. The percentiles of the Lognormal distribution are thus not difficult to calculate (Lee & Wang, 2003).

The survival and probability density are given by:

$$S(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_t^{\infty} \frac{1}{x} \exp\left[-\frac{1}{2\sigma^2}(\log x - \mu)^2\right] dx \quad (25)$$

$$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(\log x - \mu)^2\right], t > 0, \theta > 2 \quad (26)$$

Let $a = \exp(-\mu)$, then $-\mu = \log a$, and then equations (27) and (28) become

$$S(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_t^{\infty} \frac{1}{x} \exp\left[-\frac{1}{2\sigma^2}(\log ax)^2\right] dx \quad (27)$$

$$S(t) = 1 - G\left[\log\frac{at}{\sigma}\right] \quad (28)$$

$G(y)$ is the cumulative distribution function of a normal standard variable and given by the equation below:

$$G(y) = \frac{1}{\sqrt{2}} \int_0^y e^{-(u^2/2)} du \quad (29)$$

$$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(\log at)^2\right] \quad (30)$$

The Lognormal distribution is setup by the two parameters μ and 2σ . Time t cannot take on a zero value since $\log t$ is not exist for $t = 0$. The distribution is flexible and positively alligned

and the bigger the value of 2σ , the bigger the skewness (Lee & Wang, 2003).

The hazard function associated with the lognormal distribution is given by;

$$h(t) = \frac{\left(\frac{1}{t\sigma\sqrt{2\pi}}\right) \exp\left[-\frac{(\log at)^2}{2\sigma^2}\right]}{1 - G\left[\log\frac{at}{\sigma}\right]} \quad (31)$$

3.3.4 Proportional Hazard (PH) Modeling

Proportional hazard models are either parametric or semi-parametric. Parametric models aid in the interpretation of the parameters and some functions, especially the hazard rate. The semi-parametric model does not require the choice of a particular probability distribution to represent the survival time. The semi-parametric models have unspecified baseline hazard while the parametric models have the baseline hazard specified to follow a particular distribution. They assume that there is an underlying hazard rate over time, differences in the relative hazard rate at a point in time, is as a result differences in covariates. In other words, they assume no interaction between time and covariates.

Mathematically,

$$h(t) = h_0(t)e^{\beta_1x_1+\beta_2x_2+\dots+\beta_kx_k} \quad (32)$$

where $h_0(t)$ is some time function of the hazard rate known as the baseline hazard.

They are termed the proportional hazard models because the hazard for one individual is a fixed proportion of the hazard for other individual anywhere.(Klein, et al., 2003).

To see this, take the ratio of the hazards for two individuals i and j , with covariates

$$\frac{(h|Z)}{(h|Z^*)} = \frac{h_0(t)\exp\left[\sum_{k=1}^p \beta_k X_k\right]}{h_0(t)\exp\left[\sum_{k=1}^p \beta_k X_k^*\right]} = \exp\left[\sum_{k=1}^p \beta_k (X_k - X_k^*)\right] \quad (33)$$

Key to this equation is that $h_0(t)$ strikes out in both the numerator and denominator, making the ratio of the hazards to be constant over time.

Proportional hazards models have the following strong cases:

- Proportional hazards modeling utilizes information on survival time (for instance time to promotion), rather than using just a simple dependent variable that is dichotomous.
- The proportional hazards model allows us to for instance, differentiate between an employee who survived the event in year one and an employee who did same in another year. This is an imperative distinction, as studies have suggested that failure to consider the timing of employee job movements may yield biased findings (Morita, et al., 1993).
- Again, proportional hazards modeling conveniently caters for censored data. In some cases, the exact survival time is unknown, although it is known to be greater than the specified value. Censoring occurs when the study ends without all the employees getting promoted.

If the effect of censoring was ignored, the sample size and power will reduce and in effect render the results biased.

There are tests available to check data suitability for PH modeling. Tests of PH assumptions for parametric and semi-parametric tests take similar formats. If we graph the log hazards for any two individuals for instance, the proportional hazards property implies that the hazard functions should be strictly parallel (Allison, 1995).

However, to test the PH assumption, a plot of $\log(-\log S(t))$ against $\log(t)$ is an appropriate graphical test. The Survival estimates used for this plot are KM survival estimates of two or more levels of covariates and parallel and linear lines indicates that the PH assumption hence AFT assumption holds. And therefore the use of a Weibull distribution on the survival data is reasonable. If the line in the plot is straight and the slope is equal to 1, then either an AFT or PH exponential modeling is appropriate.

When the lines are parallel but not linear, a Cox Model can be used because it is PH but not Weibull hence not AFT. If the lines are straight but not parallel, then neither PH or AFT holds, but Weibull with a different shape parameter can be used. Finally, if the lines are neither straight nor parallel, the PH assumption is violated and hence the data does not follow a Weibull distribution. The points above are summarised as:

Summary of possible results for plot of $\log(-\log S(t))$ against $\log(t)$

1. Straight parallel lines \Rightarrow PH, AFT and Weibull assumptions hold
 2. Straight parallel lines with slope of 1 \Rightarrow PH, AFT and Exponential.
 3. Parallel but not straight lines \Rightarrow Not AFT, PH, but not Weibull, (can use Cox model)
 4. Non-parallel and not straight lines \Rightarrow PH violated, Not Weibull
 5. Straight lines but not parallel \Rightarrow PH and AFT violated but Weibull holds.
- Different. (Klein & Kleinbaum,2005)

The PH assumption demands that the Hazard Ratio (HR) is same over time, or equivalently, that the hazard for one person is proportional to the hazard for any other person, where the proportionality constant, which is not reliant on time, denote the set of X 's for two individuals. Hence if the test of assumption is violated for a given PH modeling technique, that method will not be executed.

In investigating a group, the length of survival is related to various characteristics. And therefore an underlying hazard for the “average” subject denoted by $h_0(t)$ is used to specify the hazard function. This is written as:

$$h(t) = h_0(t)h(t, X) \quad (34)$$

where $h(t, X)$ is a function that may change with time . Equation (36) can be rewritten as,

$$h(t, X) = \frac{h(t)}{h_0(t)} \quad (35)$$

If both $h(t)$ and $h_0(t)$ change with time, their ratio can remain constant thus a PH; this implies that, at any given time t , the hazard rate applying to a subject will be times that of an average subject for a constant $h(t, X) = h$ not changing with time.

The logarithm of hazard and of the HR are often used and written as:

$$\log[h(t, X)] = \log h = \log \left[\frac{h(t)}{h_0(t)} \right] = \beta \quad (36)$$

such that β and h do not depend on time.

The strong case for the PH model developed by Cox (1972) is not only that it permits survival data resulting from a non-constant hazard rate to be modelled, but it does so without making any assumption about the fundamental distribution of the hazards in different groups, except that the hazards in the groups remain proportional to each other over time. (Machin, Cheung & Parmar).

The PH and HR can be extended to compare two groups; one group is treated as a control or reference group $h_C(t)$ and the other group used in comparison is the treatment group $h_N(t)$

The HR for comparing two groups is derived as:

$$\begin{aligned} h_C(t) &= h_0(t) \\ h_N(t) &= h_0(t)\exp(\beta) \\ HR &= \frac{h_N(t)}{h_C(t)} = \frac{h_0(t)\exp(\beta)}{h_0(t)} \\ HR &= \exp(\beta) \end{aligned} \quad (37)$$

In PH modeling for a given coefficient say β_i , the coefficient and hazard ratios are interpreted as follows;

- a. For $\beta_i < 0$, the hazard associated with the coefficient is higher and hence the survival is better since $\exp(\beta_i) < 1$.
- b. For $\beta_i = 0$, there is no association between the hazard and coefficient because $\exp(\beta_i) = 1$.
- c. For $\beta_i > 0$, there is a higher hazard and therefore a poorer survival associated with the coefficient because $\exp(\beta_i) > 1$.

Parametric PH Models

The parametric distributions which can be modeled using the PH procedure include the Exponential, Weibull and Gompertz distributions.

While the Exponential and Weibull employ both AFT and PH modeling techniques, the Gompertz uses solely PH modeling approach to analyse survival data.

In estimating parametric survival models, time is assumed to follow some distribution whose probability density function, $f(t)$. The survival function can be modelled for both censored and non-censored data which makes it a better option for modeling as compared to the density function.

Also, the hazard function is also more interesting than the survival function because it examines an individual's risk of experiencing an event of interest as compared to the survival which looks at the experience of the event by every subject in the long run.

Plots generated from survival estimates that were obtained from parametric survival models are mostly consistent with a theoretical survival curve. If the researcher is comfortable with the fundamental distributional assumption, then parameters can be estimated that aptly specify the survival and hazard functions.

This uncomplicatedness and comprehensiveness are the main appeals of using a parametric approach (Kleinbaum & Klein, 2005).

The exponential distribution has a constant hazard of $h(t) = \lambda$ hence this is substituted into (36) as:

$$h(t, X) = \lambda \sum_{i=1}^n (\beta_i X_i) \quad (38)$$

and the log of the hazard is also deduced in order to linearize the equation, thus equation (17) becomes:

$$\log h(t, X) = \log \lambda + \sum_{i=1}^n (\beta_i X_i) \quad (39)$$

However, the hazard of the exponential function is re-parameterized as

$$h(t) = \lambda, \text{ where}$$

$$\lambda = \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k) \quad (40)$$

written as:

$$h(t, X) = \lambda(t) \exp \left[\beta_0 + \sum_{i=1}^k (\beta_i X_i) \right] \quad (41)$$

The Weibull and exponential distribution are re-parameterized and specified as the baseline hazard for the equation to become a parametric PH model.

For the Weibull with a hazard function $h(t) = \lambda \alpha t^{\alpha-1}$ is re-parameterized as

$$\lambda = \exp \left[\beta_0 + \sum_{i=1}^k (\beta_i X_i) \right] \quad (42)$$

Hence the parametric form of the equation is substituted as:

$$h(t) = \alpha t^{\alpha-1} \exp \left[\beta_0 + \sum_{i=1}^k (\beta_i X_i) \right] \times \exp \left[\beta_0 + \sum_{i=1}^k (\beta_i X_i) \right] \quad (43)$$

Thus equation (43) is the Weibull proportional hazard model for the covariates investigated under the time to first promotion of lecturers.

The exponential proportional hazard model is similar to the Weibull and shown below in equation (44) below:

$$h(t) = \exp \left[\beta_0 + \sum_{i=1}^k (\beta_i X_i) \right] \times \exp \left[\beta_0 + \sum_{i=1}^k (\beta_i X_i) \right] \quad (44)$$

for the hazard of the exponential λ re-parameterized as

$$\lambda = \exp \left[\beta_0 + \sum_{i=1}^k (\beta_i X_i) \right]$$

Again, a Gompertz PH model is obtained when the baseline hazard of the PH model is specified as a Gompertz distribution $h_0 = [\exp(\gamma t)]$ is given by:

$$h(t, X) = [\exp(\gamma t)] \times \exp(\beta_0 + \beta_1 \text{Covariate}) \quad (45)$$

3.3.5 AFT Modeling

Hazard functions are described as functions of explanatory variables but it is possible to let the explanatory variables act via a scale factor directly on time. (Hougaard, 2000).

The AFT model can be expressed on a log scale as an additive term as:

$$\log(T) = \alpha_0 + \sum_{i=1}^k \alpha_i X_i + \sigma \varepsilon_i \quad (46)$$

but multiplicative on T and given as,

$$T = \exp(\alpha_0) \cdot \exp \left(\sum_{i=1}^k \alpha_i X_i \right) \cdot \exp(\sigma \varepsilon_i) \quad (47)$$

Where α_i are the unknown parameters, X_i are the covariates of interest for $i = 1, 2, 3, \dots, k$ and ε_i is the random error of the equation which follows a particular distribution. The σ scales the error term and is re-parameterized in R as $\sigma = \frac{1}{p}$, therefore equation (46) and (47)

become:

$$\log(T) = \alpha_0 + \sum_{i=1}^k \alpha_i X_i + \frac{1}{p} \varepsilon_i \text{ and } T = \exp(\alpha_0) \cdot \exp\left(\sum_{i=1}^k \alpha_i X_i\right) \cdot \exp\left(\frac{1}{p} \varepsilon_i\right)$$

respectively.

Therefore, for the Weibull model, t is obtained by,

$$\begin{aligned} S(t) = \exp(-\lambda t^\alpha) &\Leftrightarrow -\log(S(t)) = \lambda t^\alpha \\ \Rightarrow t &= [-\log(S(t))]^{1/\alpha} \left[\frac{1}{\lambda^{1/\alpha}} \right] \end{aligned} \quad (48)$$

When it is re-parameterized,

$$\begin{aligned} \frac{1}{\lambda^{1/\alpha}} &= \exp(\alpha_0 + \alpha_1 \text{Covariate}) \\ t &= [-\log(S(t))]^{1/\alpha} \exp(\alpha_0 + \alpha_1 \text{Covariate}) \end{aligned} \quad (49)$$

Hence the acceleration factor for a fixed value $S(t) = q$ is calculated for instance for Gender with levels *Female* = 0 and *Male* = 1, the acceleration factor γ is

$$\begin{aligned} \gamma &= \frac{[-\log(q)]^{1/\alpha} \exp(\alpha_0 + \alpha_1(1))}{[-\log(q)]^{1/\alpha} \exp(\alpha_0 + \alpha_1(0))} \\ \gamma &= \exp(\alpha_0 + \alpha_1(1) - \alpha_0) \\ \gamma &= \exp(\alpha_1) \end{aligned} \quad (50)$$

The exponential model like the Weibull is re-parameterized in similar fashion.

$$\begin{aligned} S(t) &= \exp(-\lambda t) \\ \Rightarrow t &= [-\log(S(t))] \times \lambda^{-1} \end{aligned} \quad (51)$$

Let $\lambda^{-1} = \exp(\alpha_0 + \alpha_1 \text{Covariate})$, then equation (53) becomes

$$t = [-\log(S(t))] \times \exp(\alpha_0 + \alpha_1 \text{Covariate}) \quad (52)$$

Thus, the acceleration factor is obtained by;

$$\begin{aligned} \gamma &= \frac{[-\log(q)] \exp(\alpha_0 + \alpha_1(1))}{[-\log(q)] \exp(\alpha_0 + \alpha_1(0))} \\ \gamma &= \exp(\alpha_0 + \alpha_1(1) - \alpha_0) \\ \gamma &= \exp(\alpha_1) \end{aligned} \quad (53)$$

for a covariate with two levels 0 and 1 and $S(t) = q$.

The survival function of the log-logistic model is given by,

$$S(t) = \frac{1}{1 + \lambda t^p}$$

and can be rewritten as

$$S(t) = \frac{1}{1 + (\lambda^{1/p} t)^p}$$

The equation is written in terms of t as

$$t = \left[\frac{1}{S(t)} - 1 \right]^{1/p} \times (\lambda^{1/p})^{-1}$$

where the shape parameter is re-parameterized as

$$\lambda^{1/p} = \exp(\alpha_0 + \alpha_1 \text{Covariate}).$$

The acceleration factor is deduced for $S(t) = q$ as:

$$t = [q^{-1} - 1]^{1/p} \times \exp(\alpha_0 + \alpha_1 \text{Covariate}) \quad (54)$$

Therefore for a dummy variable with two levels, the acceleration factor for the log-logistic model is;

$$\gamma = \frac{[q^{-1} - 1]^{1/p} \exp(\alpha_0 + \alpha_1(1))}{[q^{-1} - 1]^{1/p} \exp(\alpha_0 + \alpha_1(0))}$$

$$\gamma = \exp(\alpha_0 + \alpha_1(1) - \alpha_0)$$

$$\gamma = \exp(\alpha_1) \quad (54)$$

3.3.6 Diagnostic Tests

In order to use PH and AFT models, it is prudent to check for data suitability. There various methods of achieving this aim, some of which are discussed below and employed in this work. Proportional hazard models are tested for suitability by using analytical or graphical tests. The test is based on chi-square tests or large sample Z tests are used to assess whether the hazard ratio of two observations is constant. Each covariate is tested individually, for a p-value, say greater than 0.10, the PH assumption is reasonable, whereas a small p-value, say less than 0.05, suggests that the variable being tested does not satisfy this assumption. The analytical test is a more powerful test because decisions are based on p-value unlike the graphical approach which is concluded based on a researcher's observation.

Plots available in using the graphical approach include:

- (a) Plots of survival estimates for two groups, using KM estimates
- (b) Plots of log negative log estimated survival function versus log of time for two or more subgroups, using KM estimates
- (c) Plots of observed survival probabilities versus expected under PH model
- (d) Plots of weighted Schoenfeld residuals from Cox models verses time Schoenfeld residuals are stated for every subject who has an event, for each independent variable in the model.

To interpret plots of log negative log estimated survival function versus log of time for two or more subgroups, using KM estimates for categories, parallel lines indicate that PH assumptions hold. In a situation where the PH assumptions do not hold, the covariates interaction with time is suspected. However, there are some interventions that can make the data suitable or the methodology in analysis is switched. Some of the interventions are:

1. Use a parametric model whose PH assumption is not very relevant.
2. Examine the physical form of the covariates in the model. They might have a more complex relationship with the log hazard.
3. Perform a repeat of the analysis by stratifying on the independent variables; if there are no other covariates of interest then do not fit any model, just obtain Kaplan–Meier curves for each covariate group separately. If there are additional control variables, fit a Cox model in which the levels of the independent variable are treated as strata.
4. Begin the analysis at a time when the PH assumption appears to be unviolated, and use a Cox PH model for only those individuals that survived that long.
5. Obtain two different hazard ratio estimates by fitting one Cox model to the early data and a different Cox model for the later data, one for each of the two time periods.
6. Fit a improved Cox model that includes a time-dependent variable that integrates the interaction of independence with time. Such a model is called an extended Cox model. (Weiss, 2010).

There are available tests in order to ascertain whether the models have been correctly chosen, these tests are referred to as goodness-of-fit tests. There are tests of subsets of parameters in a distribution, all parameters in a distribution, appropriateness of distribution family, among several other tests.

To pick an appropriate parametric model for survival data, the generalized gamma model is commonly used, but rarely for a final one. The Likelihood or Wald ratio each tests the hypotheses that $\theta = 1$ or $\theta = 0$ and provides a means of checking the assumption of a Weibull or Log normal regression model respectively. It is difficult to use a formal statistical test to discriminate between parametric models because the models are not nested in a larger model which includes all the regression models discussed in this chapter. It is only the Weibull and Log normal distribution that defy this.

One way of choosing an appropriate parametric model is to base the decision on the value Akaike Information Criterion (AIC). The AIC is a vital statistic in selecting the best-fitting model in the set. It must however be noted that, if all the models are very poor, AIC will still select the one estimated to be best, even though relatively fitting model might be in itself poor in an absolute sense. Thus, every effort must be made to ensure that the set of models is well founded (Burnham & Anderson). For the parametric models talked about above, the AIC is given by

$$AIC = -2 * \log(Likelihood) + 2(p + k)$$

where p is the number of regression parameters and k is the number of parameters (Klein & Moeschberger, 2003).

The Wald Test, Log-Likelihood and Score tests however, are tests of parameters, while the AIC and Bayesian Information Criterion (BIC) are also procedures for selecting best-fitting model.

3.3.7 Modeling and Model Selection

As discussed above, there are two main modeling techniques in survival analysis; PH and AFT modeling and both approaches are employed in this work.

The AIC values of the models were used to choose the best fitting model. The best model ideally, should have the smallest AIC as possible. The AIC values of parametric models for the PH model are compared separately and then compared to the semi-parametric model.

Then, the best-fitting AFT model is also selected based on the AIC values.

3.4 Statistical Software

There are various statistical software available to analyse survival data, these include SPSS, SAS, STATA, and R; to mention a few. In this research however, the data was obtained in the form of a MS excel spreadsheet.

Analyses were widely done using STATA. The AFT and PH models were built using STATA with the package *stgenreg*.

3.5 Dissemination of Results

A copy of this dissertation will be placed in University of Ghana library.

CHAPTER FOUR

RESULTS AND DISCUSSION

This chapter presents the results, interpretation and discussion of the study. Preceding the analysis are the general characteristics of the data, variables coding and interpretation. The analysis starts with a descriptive statistics with the application of some non-parametric statistical models and the final analysis is based on the selection criteria of the best models and some useful results from the best selected model.

4.1 Explanatory Variables

This study considers the variables that have an effect on time until first promotion of a lecturer which is the response variable of interest. The variables include socio demographic variables consisting of sex, age, number of children and marital status. The others are the academic requirement variables which include number of UG committees a lecturer has membership, number of national committees a lecturer has membership, international conferences attended, books written, technical reports written, number of working papers, country of origin of lecturer's certificate, qualification before entry, lecturer's college of affiliation. The above-mentioned variables can also be grouped into continuous and categorical variables. The continuous variables include age, number of children, number of UG committees one has membership, number of national committees one has membership, international conferences attended, books written, technical reports written and number of working papers.

The categorical variables include sex, marital status, qualification before entry, origin of lecturer's certificate and the college of affiliation.

Table 4.1 Variable Coding for Categorical Variables

Variables	Description	Categories
Local/Foreign	Source of lecturer's	Local (obtained Master's degree locally)=1
	postgraduate degree	Foreign (obtained Master's degree outside)=2
Qualification	Qualification before	MSc/MA/MPhil/MPH=1
	entry as lecturer	PHD=2
Marital Status	Marital status of the	Single=1
	lecturer	Married=2
Sex	Sex of the lecturer	Male =1 and Female = 2
College	College where a lecturer	CBAS- College of Basic and Applied Sciences
	lectures at the	COE- College of Education
	University.	CHS- College of Health Sciences
		COH- College of Humanities

The *event* has been coded below as:

$$Event = \begin{cases} 1 & \text{Lecturer Promoted} \\ 0 & \text{Lecturer Censored} \end{cases}$$

4.2 Preliminary Analysis

The data is a primary data collected from 151 full time lecturers from the University of Ghana using a structured questionnaire. From Table 4.2, the lecturers consist of 91 males out of which 38 of them got promoted and 53 were censored (had not been promoted). There are 60 females lecturers who responded 15 of which were promoted and 45 were censored.

The *Pearson's Chi-Square* obtained [$\chi^2 = 4.4578, p - value = 0.035$] suggests that there is a significant difference in the promotion history of males and females.

As shown in Figure 4.7 found at the Appendix section of this thesis, 33 lecturers who responded are from the College of Basic and Applied Sciences (CBAS) where only 3 had been promoted leaving 30 censored. Ten lecturers responded from the College of Education (COE) out of which 7 had been promoted and 3 were censored. Out of 59 lecturers from the College of Health Sciences (CHS), 24 had earned promotion and 35 censored. There was a total of 49 lecturers from the College of Humanities (COH), 19 of which had experienced the event of interest and 30 had not. The *Pearson's Chi-Square* reveals that there is a significant difference in the survival times of the lectures from the different colleges [$\chi^2 = 16.2430$ $p - value = 0.001$].

The country where a lecturer had their Master's degree was categorized into two groups, namely *Local degree* for lecturers who had their Master's degree in Ghana and *Foreign degree* for the lecturers who had their Master's degree outside Ghana. From Table 4.2, out of the 98 lecturers who had their Master's degree in Ghana, 23 of them were promoted. The total number of respondents with foreign Master's degree was 53, out of which 30 had been promoted. The *Pearson's Chi-Square* and the p-value revealed a significant difference between these two groups [$\chi^2 = 16.576$, $p - value = 0.001$].

Table 4.2 also presents results of lecturers' qualification before assuming post as a lecturer at the University of Ghana. From the results, 70 of the lecturers entered the University's teaching staff with MSc/MA/MPhil degree of which 25 had been promoted as at the time of collecting data for this research and 45 were censored. There were 81 lecturers who entered with PHD out of which 28 had been promoted. The *Pearson's Chi-Square* shows no significant difference between the promotion histories of the two groups at 0.05 significant level as evidenced by the result. [$\chi^2 = 0.0217$, $p - value = 0.883$],

Finally, for the marital status, there were 20 single lectures, out of which 3 had been promoted. Also, there were as 132 married lecturers with 51 having been promoted and 81 censored. The Pearson's Chi-Square and the p-values show a significant difference between the two groups [$\chi^2 = 5.7614$, $p - value = 0.016$]

Table 4.2 Variable Description for Categorical Variables

Variable	N	Promoted	Censored	Min	Max	Mean	Std. Deviation	χ^2	P-value
Gender				1	2	1.3974	0.0399	4.4578	0.035
Male	91	38	53						
Female	60	15	45						
College				1	4	2.8212	1.1140	16.243	0.001
CBAS	33	3	30						
COE	10	7	3						
CHS	59	24	35						
COH	49	19	30						
Master's Country				1	2	1.3501	0.4788	16.576	0.000
Local Master's deg.	98	23	75						
Foreign Masters deg	53	30	23						
Qual. Before entry				1	2	1.5364	0.5003	0.0217	0.883
Msc/MA/MPhil	70	25	45						
PHD	81	28	53						
Marital Status				1	2	1.8743	0.3327	5.7614	0.016
Single	20	3	17						
Married	132	51	81						

Additionally, Table 4.3 presents a summary of the continuous variables of interest in this study. For the variable age, the average age of the lecturers is 42.51 years with a standard deviation of 8.27. The minimum age of a lecturer was 28 years and the maximum age of a lecturer is 66years. The average number of children born to lecturers was 2.325 with 95% Confidence Interval (CI) of [2.053, 2.596]. The number of UG committee membership recorded an average of 3.07 with 95% CI of [2.401, 3.745]. The average number of national committee membership

is 1.252 with the maximum number recorded as 14 committees. On the number of international conferences attended by these university lecturers, the average number recorded was 10.079 where the maximum recorded was 44. Considering technical report, the average number of technical reports by respondents recorded was 0.974 with 95% CI of [0.553, 1.394]. On the number of books published by the lecturers, the table shows an average of 0.437 with a maximum number of books published as 4. Finally, the number of working papers for the lecturers recorded an average of 0.649.

Table 4.3 - Continuous variable descriptive.

Cont. Variable	Mean	Std dev.	Min.	Max.	95% conf. interval	
Age	42.510	8.2655	28	66	41.181	43.839
Children	2.325	1.6913	0	10	2.053	2.596
UG Committees	3.073	4.1779	0	30	2.401	3.745
Nat. Committees	1.252	2.6387	0	14	0.827	1.676
Int'l conferences att.	10.079	9.2668	0	44	8.590	11.570
Tech. reports written	0.974	2.6177	0	12	0.553	1.394
Books published	0.437	0.8913	0	4	0.294	0.580
Working Paper	0.649	1.5240	0	6	0.404	0.894

The Kaplan Meier survival analysis also provides a descriptive of the nature of the distribution. It provides an exploration of the survival time history of respondents. The average survival time for promotion is 78.847 months with a standard deviation of 3.72 and a 95% CI [71.497, 86.198]. This means it take an average of 6.57 years for a lecturer to be promoted.

The Kaplan Meier analysis is used to study differences which exist between groups of variables by providing a pictorial view of the patterns of survival. The group whose survival curve lies above the other is said to have a higher survival than the other defined by a lower curve. The

Kaplan Meier curves for the covariates *Gender*, *Marital Status*, *College of Affiliation*, *Qualification Before Entry* and *Country of Master's Degree Program* have been provided in Figures 4.2, 4.3, 4.4, 4.5, and 4.6 respectively (See Appendix A).

Let us begin with Figure 4.2 which shows that for the period between 0 to 90 months, males have a higher probability of being promoted as compared to females but this reverses after the 90th to 160th months.

We move now to Figure 4.3 which depicts that at the early period below 50 months, married lecturers are more likely to be promoted but between the 50th to the 170th month, the single lecturers have a higher probability of being promoted as compared to the married ones.

In Figure 4.4, we are presented with the Kaplan Meier curve comparing lecturer's qualification before entry. It shows that lecturers who entered with PHD are more likely to be promoted as compared to their counterparts who entered with MA/MSc/MPhil.

Now, in Figure 4.5, Kaplan Meier survival curve for the colleges where a lecturer works is presented. The nature of the curve, it must be said, does not provide conclusive evidence as to which colleges are most likely to be promoted.

A look at Figure 4.6 presents us with the curve for country where a lecturer received their Master's degree. From the curve, lecturers who obtained their Master's outside Ghana are more likely to be promoted than those who had their Master's degree locally, since the curve for the latter lies below that of the foreign degree holders. The KM plots suggest that there exist some differences among the categorical variables in each group. Hence there is the need to

investigate whether these differences are significant with the application of the *Log Rank Test* or *Wilcoxon Test* which will present the results based on the hypothesis that:

H_0 : The promotion times of lecturers among the groups are not different

H_1 : The promotion times of lecturers among the groups are different

The Log Rank and Wilcoxon tests are used because they provide an efficient method for comparing two or more survival curves where there exist some pattern of differences and where some of the observations may be censored and as well where the overall grouping may be stratified. The results are presented in Table 4.4. The table shows the Chi-square and P-values for both the Log Rank and the Wilcoxon test with a decision on the acceptance of H_0 and H_1 stated earlier.

Table 4.4 Log Rank and Wilcoxon (Breslow) test for equality of survivor functions

Variables	Log Rank		Wilcoxon		Decision
	χ^2	p-value	χ^2	p-value	
Gender	1.35	0.2445	2.10	0.1476	Fail to reject H_0
College	7.48	0.0581	5.68	0.1284	Fail to reject H_0
Local/Foreign Master's	12.24	0.0005	9.61	0.0019	Reject H_0
Qual. Before Entry (Master's/PHD)	11.75	0.0006	11.75	0.0006	Reject H_0
Marital Status	0.10	0.7532	0.67	0.4143	Fail to reject H_0

Table 4.4 above presents results relating to the significance of differences in the promotion experiences among the sub-divisions of the categorical variables. The table also presents the decision on whether the hypothesis of “*no difference*” should be accepted or not.

On the covariate *Gender*, there is no significant difference between the promotion of females and that of males and therefore we fail to reject the hypothesis of no difference.

This decision is confirmed by the *Chi-Square* and the *p-values* of both the Log Rank test and the Wilcoxon test from Table 4.4. On the variable *College*, the values of Chi-Square and *p-value* from the Log Rank test and the Wilcoxon test suggest that there is no significant difference in the promotion of lecturers from different colleges. From the different categories of *Marital Status*, the table reveals that there is no significant difference in the promotion of both single and married lecturers.

However, the Chi-Square test shows that there is a significant difference in the promotion lecturers who obtained their Master's degree locally and those who obtained it outside Ghana. This is in line with the KM plot which seems to project that foreign Master's degree holders earn promotion earlier as compared to the local Master's degree holders. From the same table, *Lecturers Qualification* before entering into lectureship was also significant. Thus, there is a significant difference in the promotion history of lecturers who entered into teaching with Master's degree and those who entered with a PhD. As such, the claim that PhD holders have a better chance at promotion than Master's degree holders by the KM plot is confirmed by the Log Rank test and the Wilcoxon test as being significant.

4.3 Modeling Process and Analysis

In this section, the study verifies some underlying assumptions surrounding some of the models used. The section will also compare the performance of the models based on their log-likelihood and AIC values. This rigorous process will result in we coming out with the best predicting model.

Checking the Proportional Hazard (PH) Assumption

The basic pre-requisite of PH model is to verify if the proportional hazard assumption is met to warrant its use. The Proportional Hazard assumption states that the hazard of any variables must be constant throughout the study period. That is, hazard of individual variables should not fluctuate (vary) with time. The test with application of *Schoenfeld Residuals* present a Global test on the overall model and also for the individual variables in the model. Variables which are statistically insignificant are said to have met the PH assumption and those which are found to be statistically significant have violated the PH assumption. From table 4.5, all the variables are insignificant at 0.05 significance level except for the variable *Marital Status* and *Number of Children* which recorded [$\chi^2 = 0.43, p - value = 0.0077$] and [$\chi^2 = 0.42, p - value = 0.0405$] respectively.

These values suggest that the variable marital status violates the PH assumption. The variables *Gender, Age, Country of Master's Degree Program, Qualification Before Entry, Number of Books, Technical Reports, UG Committees Membership, National Committees Membership, Number of International Conferences Attended, Working Paper and College of Affiliation* were all statistically insignificant. This shows that all these variables have satisfied the proportionality assumption.

The Global test which encompasses all the variables in a single test and value suggests the variables when put in a single model will not violate the PH assumption.

Table 4.5 Proportional hazard assumption test for the covariates

Covariate	rho	Chi-Square	p-value
Age	0.09469	0.43	0.5129
Marital Status	0.33273	7.11	0.0077
Children	-0.35199	4.20	0.0405
Master's Degree country	-0.15367	2.15	0.1421
Qualification before entry	-0.09823	0.76	0.3841
UG committee membership	-0.10942	0.90	0.3440
National Committee membership	0.11786	0.90	0.3427
International Conference attended	-0.00996	0.01	0.9306
Books	-0.13655	1.25	0.2638
Technical report	-0.06863	0.21	0.6476
Working paper	-0.01534	0.01	0.9095
Gender	0.08045	0.38	0.5390
College	-0.20964	2.51	0.1133
Global test	NA	19.40	0.1112

Selecting the Best Fitted Model

Before a decision is taken on the best fitted model for the *time until a lecturer is promoted* data, the Log Likelihood Ratio value is computed where the AIC values of each of the models is calculated with the application of the log likelihood ratio. The decision is based on the principle that the model with the lowest AIC value is the best model for the data.

A look at Table 4.6 shows the log likelihood and the AIC values for the *Cox Proportional Hazard* model, *Exponential Proportional Hazard* model, *Weibull Proportional Hazard* model and *Gompertz Proportional Hazard* model.

From the table, the Cox PH has the highest AIC absolute value making it the worst performing model in modeling a lecturer's time to promotion data. The Weibull PH recorded the lowest AIC absolute value among all the PH models and hence the best model to predict a lecturer's promotion.

Table 4.6 Comparing PH models using Log-likelihood and AIC

Model	Log-likelihood(null)	Log-likelihood (model)	df	AIC
Cox PH	-202.1162	-162.1087	15	354.2175
Exponential PH	-104.5423	-82.1796	16	196.3591
Weibull PH	-77.0982	-34.1556	17	102.3112
Gompertz PH	-83.8720	-39.3321	17	112.6643

Also, Table 4.7 compares the log-likelihood and the AIC values as well but this time for the Accelerated Failure Time (AFT) models which are the *Exponential AFT*, *Weibull AFT*, *Log-logistic AFT* and *Log-normal AFT*.

From the table, the Exponential AFT recorded the highest AIC value of 196.3591 making it the least performing model in predicting the time to a lecturers promotion given the underlying factors. The Weibull AFT recorded the lowest AIC values of 102.3112 making it the best model among the AFT models. This shows that the distribution of the data is Weibull distribution since in both the PH and AFT the Weibull distribution was the best selected model.

Table 4.7 Comparing AFT models using Log-likelihood and AIC

Model	Log-likelihood(null)	Log-likelihood (model)	df	AIC
Exponential AFT	-104.5423	-82.1796	16	196.3591
Weibull AFT	-77.09816	-34.1556	17	102.3112
Lognormal AFT	-75.97775	-40.8231	17	115.6462
Log logistic AFT	-75.38348	-38.6617	17	111.3234

Table 4.8 Standard errors of Exponential, Weibull, Log-logistic and Lognormal AFT models

Variable	Category	Exp.	Weibull	Log- logistic	Log- normal
Age of lecturer	Age	0.0499	0.0127	0.0140	0.0168
Number of Children	No. of Children	0.1535	0.0553	0.0450	0.0436
UG committee membership	No. UG committee	0.0524	0.0147	0.0156	0.0178
Nat. committee membership	No. Nat. committee	0.0664	0.0170	0.0255	0.0257
International conferences	No. Int. conferences	0.0196	0.0050	0.0053	0.0064
Books published	No. Books published	0.1944	0.0526	0.0626	0.0685
Technical reports written	No. Technical reports	0.0829	0.0216	0.0252	0.0295
Working papers	No. Working Papers	0.3035	0.0923	0.0885	0.0832
Gender of Lecturer	Male	0.4247	0.1177	0.1296	0.1404
College of Lecturer	CBAS	0.6696	0.1715	0.1837	0.2127
	COE	0.5972	0.1692	0.1963	0.2365
	CHS	0.4099	0.1144	0.1224	0.1148
Local/Foreign degree	Local Master Deg.	0.3735	0.1111	0.1164	0.1212
Qualification in entry	Master Cert.	0.3758	0.1113	0.1137	0.1109
Marital Status	Married	1.0146	0.3204	0.5411	0.3835
	Constant	1.7460	0.4605	0.6625	0.6167

Equally important is Table 4.8 which compares the standard errors of the variables associated with the AFT models. It also gives the standard error for the parameter estimate for the full model. From the table it can be seen that the standard errors for the variables under the Weibull AFT are smaller as compared to the other AFT models. The standard error for the Weibull

Parameter estimate also recorded the least value. This reveals that the parameter estimation with the Weibull AFT is accurate as compared to the others.

In the end, the data is modeled with the Weibull PH model though the Weibull AFT was also adjudged the best model among the AFT models. The Weibull PH is used because it provides a clearer and simpler interpretation.

Table 4.9 as well presents the covariates, their respective hazard ratios, standard errors, the Z-score, P-value and the confidence interval for the hazard ratio. From the table, variables like the number of children by a lecturer, number of national committee membership of a lecturer, number of books published by a lecturer, marital status, college of affiliation of a lecturer and gender were not significant at 5% significant level on the time until promotion of a lecturer. But the variables Age, number of UG committee membership, number of international conferences attended, number of technical reports written, number of working papers, country of Master's degree and qualification before entry into lectureship were all significant at 5% significant level.

The results of the fitted final model in Table 4.9 are interpreted in terms of hazard ratios (HR), the logarithm of which gives the coefficients in the model. For categorical variables, they are interpreted by comparing the reference category with others. Similarly, the coefficient for a continuous explanatory variable indicates the estimated change in the logarithm of the hazard ratio for a unit increase in the value of the covariate when the remaining covariates in the model are constant. From the results presented by the Weibull PH model, the Weibull model in Table 4.9 shows that for the covariate *Age*, a unit increase in the age of a lecturer corresponds to a

decrease in the promotion time since the hazard ratio is less than 1 [$HR=0.7532$, p -value= 0.000].

For the *number of University of Ghana committee membership*, a unit increase in the number of committees is likely to increase the chances of being promoted by 18.95% [$HR = 1.1895$ p -value = 0.007].

On the *number of international conferences attended*, the table shows that a unit increase in the number of international conferences attended is likely to increase the chances of promotion by 8.43% [$HR = 1.0843$ p -value = 0.001].

The table also reveals that for a unit increase in *the number of technical reports* written by a lecturer, it is likely to increase the chances of promotion by 35.13% [$HR = 1.3513$ p -value= 0.002].

On the number of working papers possessed by the lecturer, an increase in the number is likely to decrease the chances of getting a promotion [$HR = 0.1634$ p -value= 0.000].

For the *country of lecturers Master's degree*, the variable foreign Master's was used as a reference category and it can be shown from the table those who possess a local Master's degree are less likely to be promoted compared to those with foreign degree [$HR = 0.3351$ p -value= 0.017].

Finally for the variable *qualification before entry*, PhD was used as a reference category. The value obtained from the HR suggest that lecturers who started with a PhD degree are more likely to be promoted earlier than those with a Master's degree [$HR = 0.0833$ p -value= 0.000].

From Table 4.9, it is safe to say that the model containing all the predictor variables is statistically significant with p-value less than 0.05. This indicates that the model fits well and is able to predict the probability of a lecturer being promoted earlier or not.

Table 4.9: Weibull Proportional Hazard Model

Covariates	Hazard. Ratio	Std. Err.	z	p-value	[95% Conf.	Interval]
Age	0.7532	0.0475	-4.49	0.000	0.6657	0.8524
No. of Children	1.1083	0.2577	0.44	0.658	0.7027	1.7481
No. UG committee	1.1895	0.0761	2.71	0.007	1.0490	1.3485
No. Nat. committee	1.0194	0.0731	0.27	0.788	0.8859	1.1732
No. Int. conferences	1.0843	0.0258	3.40	0.001	1.0350	1.1360
No. Books published	1.2353	0.2727	0.96	0.339	0.8014	1.9040
No. Technical reports	1.3513	0.1338	3.04	0.002	1.1129	1.6408
No. Working Papers	0.1634	0.0692	-4.28	0.000	0.0712	0.3749
Male	1.4599	0.7216	0.77	0.444	0.5541	3.8465
CBAS	0.6903	0.4961	-0.52	0.606	0.1688	2.8233
COE	1.5428	1.1039	0.61	0.545	0.3795	6.2718
CHS	5.5454	2.8386	3.35	0.001	2.0341	15.1181
Local Master Deg.	0.3351	0.1539	-2.38	0.017	0.1362	0.8245
Master Cert.	0.0833	0.0441	-4.70	0.000	0.0295	0.2351
Married	2.0369	2.7854	0.52	0.603	0.1396	29.7150
Con	0.0000	0.0002	-4.19	0.000	0.0000	0.0060
/ln_p	1.4389	0.1133	12.70	0.000	1.2168	1.6609

CHAPTER FIVE

CONCLUSION AND RECOMMENDATIONS

This final chapter presents the conclusion of the study. This final chapter will also make recommendation to aspiring lecturers on the factors that can contribute to early promotion.

5.1 Conclusions

The study first performed a descriptive test with the application of the Kaplan Meier curve where the Log Rank and the Wilcoxon test was able to reveal that among the categorical variables, qualification before entry and origin of a lecturers Master's degree was significant at 5%. This means that there exists a significant difference between the groups within these two variables. The Kaplan Meier Curve therefore showed that of the qualification before entry, lecturers who entered into teaching with a PhD and lecturers who had their Master's degree abroad were better favored in terms of promotion.

The Proportional Hazard (PH) assumption was verified with the Log Rank test which also showed that all the variables do satisfy the proportionality assumption except for the variable number of children and marital status. But the global test value was insignificant which means the full model has satisfied the PH assumption and hence a PH model can be applied to the data. The study used the Akaike Information Criterion to select the best model on the condition that the model with the least AIC value is the best model for the promotion data. The results revealed that the Weibull PH model was the best model among the PH models and Weibull AFT was the best model compared to the other AFT models.

Consequently, the Weibull PH model was chosen as the final model because it provides an easy interpretation.

The variables included in the final model were age, gender, number of children, number of UG committee membership, number of national committee membership, number international conference attended, number of technical reports written, number of working papers, marital status, number of books published, qualification before entry, origin of Master's certificate and the college of affiliation of the lecturer. The variables *number of children, marital status, gender, national committee membership, and college of affiliation* were not significant at 5% significant level.

For the variables that were significant at 5% significant level, it can be seen that the chances of getting earlier promotion decreases with an increase in age. Thus young lecturers are more likely to be promoted as compared to aged lecturers. But the higher the number of UG committee membership, International conferences attended, books published and technical reports written the more likely it is to get a promotion. For number of working papers, the higher the number the less likely you are to be promoted. For the categorical variables that were also significant at 5% significant level, it was revealed that lecturers with local Master's degree are less likely to be promoted compared to those with foreign Master's degree and finally lecturers who began lecturing with a Master's degree are less like to be promoted compared to those who started with PHD. These results suggest that there is fairness in the promotion process since the hard working lecturers usually gets promoted.

5.2 Recommendation

The results in the study show that promotion of a lecturer takes into consideration a cocktail of factors. Hence, for a lecturer to be promoted faster he/she must publish more books, attend several international conferences, be involved in university of Ghana committee activities, write more technical reports, and all these should be done at a younger age. Finally, lecturers should as well blend their degrees with foreign-acquired degrees and also work hard to acquire a PhD before lecturing.

5.3 Future Research

The study used only University of Ghana lecturers as subjects. It is suggested that for future research, lecturers from other universities be included. In future studies, lecturers' promotion over the years ahead should be observed and not a retrospective review of their promotion history which in most cases lecturers could not provide exact date for but rather gave a range. Subsequent studies should capture exact date to provide more accurate results.

REFERENCES

- Aalen, O. O. (1988). Heterogeneity in survival analysis. *Statistics in medicine*, 7(11), 1121-1137.
- Althausen, R. P., & Kalleberg, A. L. (1981). Firms, occupations, and the structure of labor markets: A conceptual analysis. *Sociological perspectives on labor markets*, 8, 119-149.
- Ambrose, M. L., & Cropanzano, R. (2003). A longitudinal analysis of organisational fairness: An examination of reactions to tenure and promotion decisions. *Journal of Applied Psychology*, 266.
- Awang, Z., Ahmad, J. H., & Zin, N. M. (2010). Modelling job satisfaction and work commitment among lecturers: A case of UiTM Kelantan. *Journal of Statistical Modeling and Analytics*, 45-59.
- Ayee, J. R. (2001). Civil service reform in Ghana: A case study of contemporary reform problems in Africa. *African Journal of Political Science/Revue Africaine de Science Politique*, 1-41.
- Barone-Adesi, G., Engle, R. F., & Mancini, L. (2008). A GARCH option pricing model with filtered historical simulation. *The review of financial studies*, 1223-1258.
- Bryant, T. (2002). Role of knowledge in public health and health promotion policy change. *Health Promotion International*, 89-98.
- Burnham, K. P., & Anderson, D. R. Information and likelihood theory: a basis for model selection and inference. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 49-97.
- Chen, Z., Veiga, J. F., & Powell, G. N. (2011). A survival analysis of the impact of boundary crossings on managerial career advancement up to midcareer. *Journal of Vocational Behavior*, 79(1), 230-240.

- Enders, J. (2001). A chair system in transition: Appointments, promotions, and gate-keeping in German higher education. *Higher education*, 3-25.
- Forster, N. (2001). A case study of women academics' views on equal opportunities, career prospects and work-family conflicts in a UK university. *Career Development International*, 6(1), 28-38.
- Giudici, P. (2005). *Applied data mining: statistical methods for business and industry*. John Wiley & Sons.
- Gruneberg, M. M., & Startup, R. (1978). The job satisfaction of university teachers. *The Vocational Aspect of Education*, 30(76), 75-79.
- Gruneberg, M. M., Startup, R., & Tapsfield, P. (1974). A Study of University Teachers' Satisfaction with promotion procedures. *The Vocational Aspect of Education*, 26(64), 53-57.
- Gutierrez, R. G. (2002). Parametric frailty and shared frailty survival models. *Stata Journal*, 2(1), 22-44.
- Henderson, R., & Oman, P. (1999). Effect of frailty on marginal regression estimates in survival analysis. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 61(2), 367-379.
- Hosmer, D. W., Lemeshow, S., & May, S. (1999). Descriptive methods for survival data. *Applied Survival Analysis: Regression Modeling of Time-to-Event Data, Second Edition*, 16-66.
- Hougaard, P. (2000). Univariate Survival Data. *Analysis of Multivariate Survival Data*, 36-111.
- Johnson, G. E., & Stafford, F. P. (1974). The earnings and promotion of women faculty. *The American Economic Review*, 64(6), 888-903.

- Johnston, R. J., & Brack, E. V. (1983). Appointment and promotion in the academic labour market: a preliminary survey of British University Departments of Geography, 1933-1982. *Transactions of the Institute of British Geographers*, 100-111.
- Karim, N. H. A. (1970). Investigating the correlates and predictors of job satisfaction among Malaysian academic librarians. *Journal for Information Systems Research and Practice*, 13(2).
- Khalid, S., Irshad, M. Z., & Mahmood, B. (2012). Job satisfaction among academic staff: A comparative analysis between public and private sector universities of Punjab, Pakistan. *International Journal of Business and Management*, 7(1), 126.
- Klein, J. P., & Moeschberger, M. L. (2005). *Survival analysis: techniques for censored and truncated data*. Springer Science & Business Media.
- Lee, E. T., & Wang, J. W. (2003). Some Well-Known Parametric Survival Distributions and Their Applications. *Statistical Methods for Survival Data Analysis, Third Edition*, 134-161.
- Locke, E. A., Fitzpatrick, W., & White, F. M. (1983). Job satisfaction and role clarity among university and college faculty. *The Review of Higher Education*, 6(4), 343-365.
- Macfarlane, B. (2007). Defining and rewarding academic citizenship: The implications for university promotions policy. *Journal of Higher Education Policy and Management*, 29(3), 261-273.
- Machin, D., Cheung, Y. B., & Parmar, M. K. (2006). Cox's proportional hazards model. *Survival Analysis: A Practical Approach, Second Edition*, 121-153.
- Malik, M. E., Nawab, S., Naeem, B., & Danish, R. Q. (2010). Job satisfaction and organisational commitment of university teachers in public sector of Pakistan. *International Journal of Business and Management*, 5(6), 17.

- Martinez, D. R., Cabaleiro, J. C., Pena, T. F., Rivera, F. F., & Perez, V. B. (2010, February). Performance modeling of MPI applications using model selection techniques. In *Parallel, Distributed and Network-Based Processing (PDP), 2010 18th Euromicro International Conference on IEEE*.
- Morita, J. G., Lee, T. W., & Mowday, R. T. (1993). The regression-analog to survival analysis: A selected application to turnover research. *Academy of Management Journal*, 36(6), 1430-1464.
- Mottaz, C. J. (1985). The relative importance of intrinsic and extrinsic rewards as determinants of work satisfaction. *The Sociological Quarterly*, 26(3), 365-385.
- Oshagbemi, T. (2000). Gender differences in the job satisfaction of university teachers. *Women in Management review*, 15(7), 331-343.
- Özbilgin, M., & Healy, G. (2004). The gendered nature of career development of university professors: The case of Turkey. *Journal of Vocational Behavior*, 64(2), 358-371.
- Toker, B. (2011). Job satisfaction of academic staff: an empirical study on Turkey. *Quality Assurance in Education*, 19(2), 156-169.
- Singh, R., & Mukhopadhyay, K. (2011). Survival analysis in clinical trials: Basics and must know areas. *Perspectives in clinical research*, 2(4), 145.
- Lawless, J. F. (2011). *Statistical models and methods for lifetime data* (Vol. 362). John Wiley & Sons.
- Fernandez, J. P. (1981). *Racism and sexism in corporate life: Changing values in American business*. Free Press.

APPENDIX A

Figure 4.1 Kaplan-Meier Curve

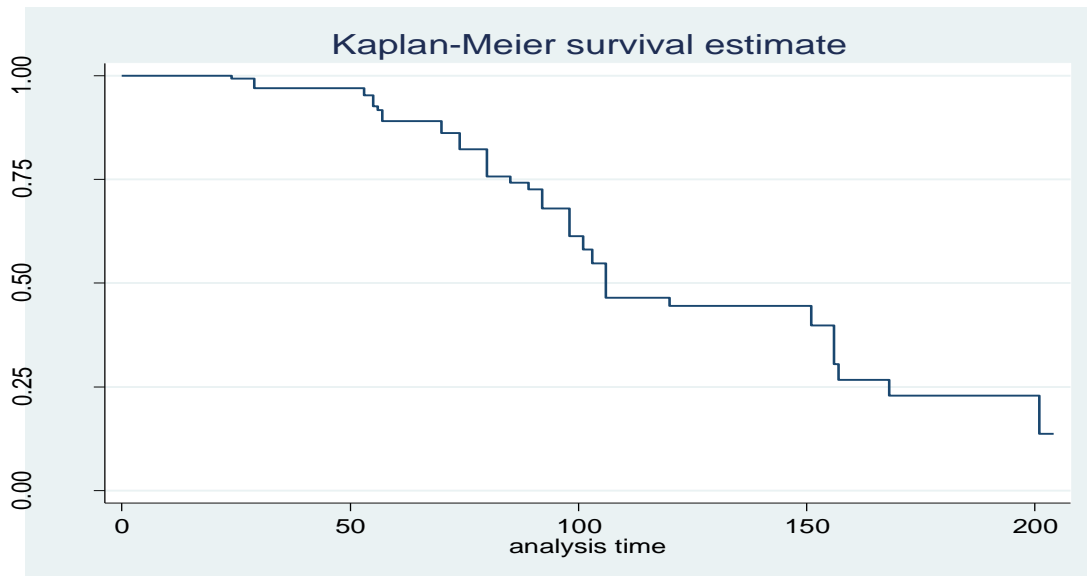


Figure 4.2 Kaplan-Meier Curve comparing Gender

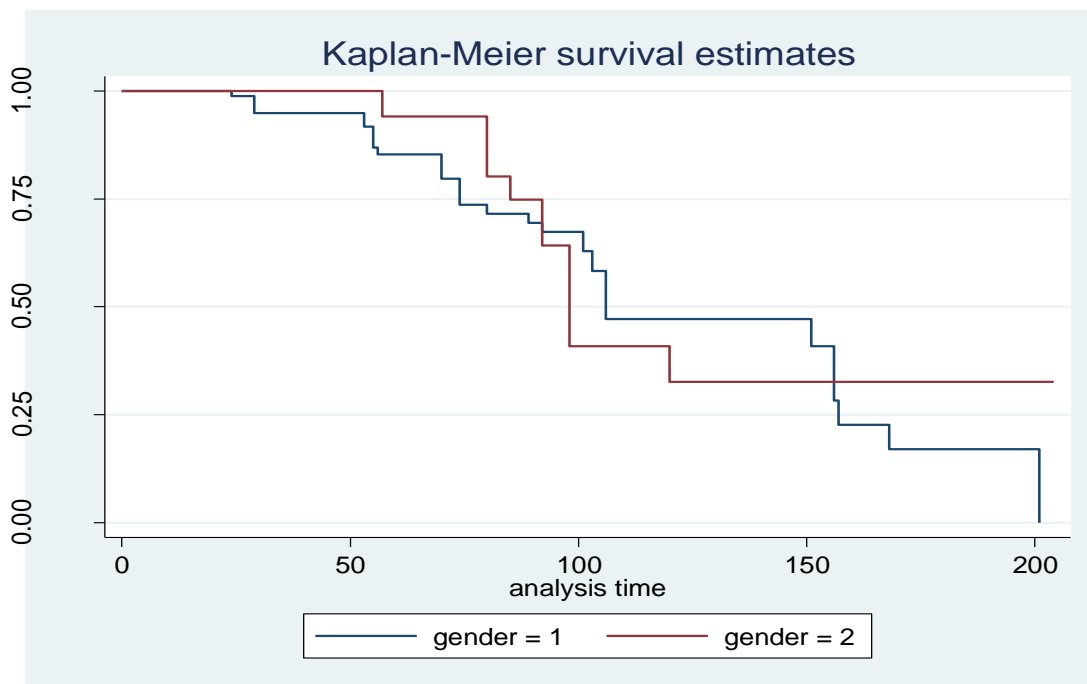


Figure 4.3 Kaplan-Meier Curve comparing Colleges

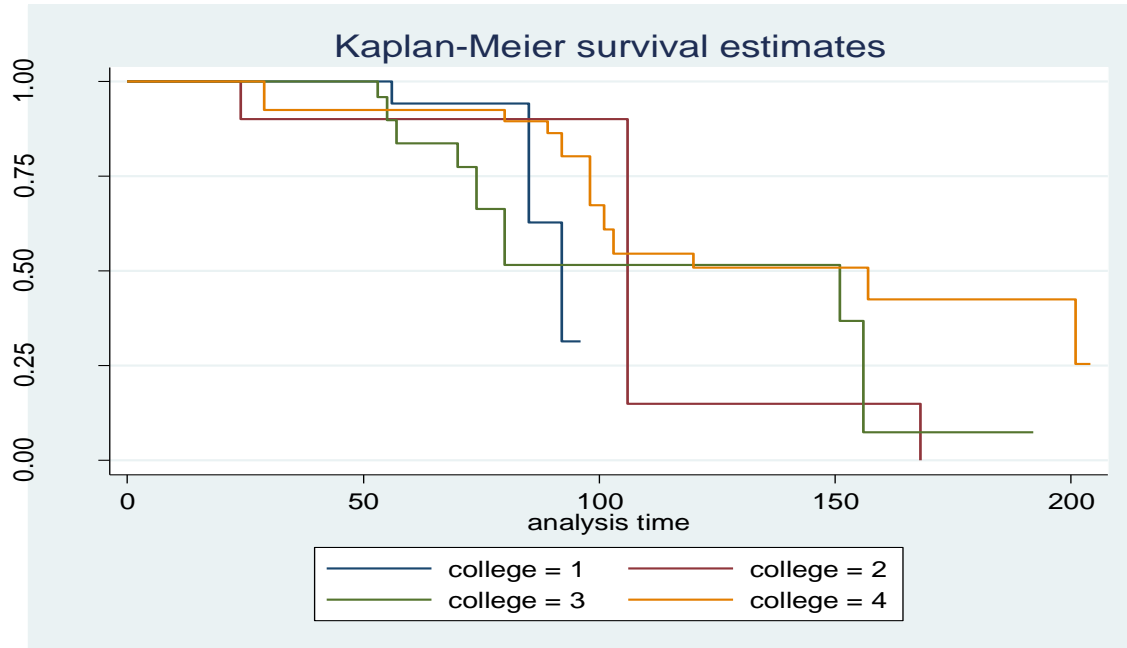


Figure 4.4 Kaplan-Meier Curve comparing Marital Status

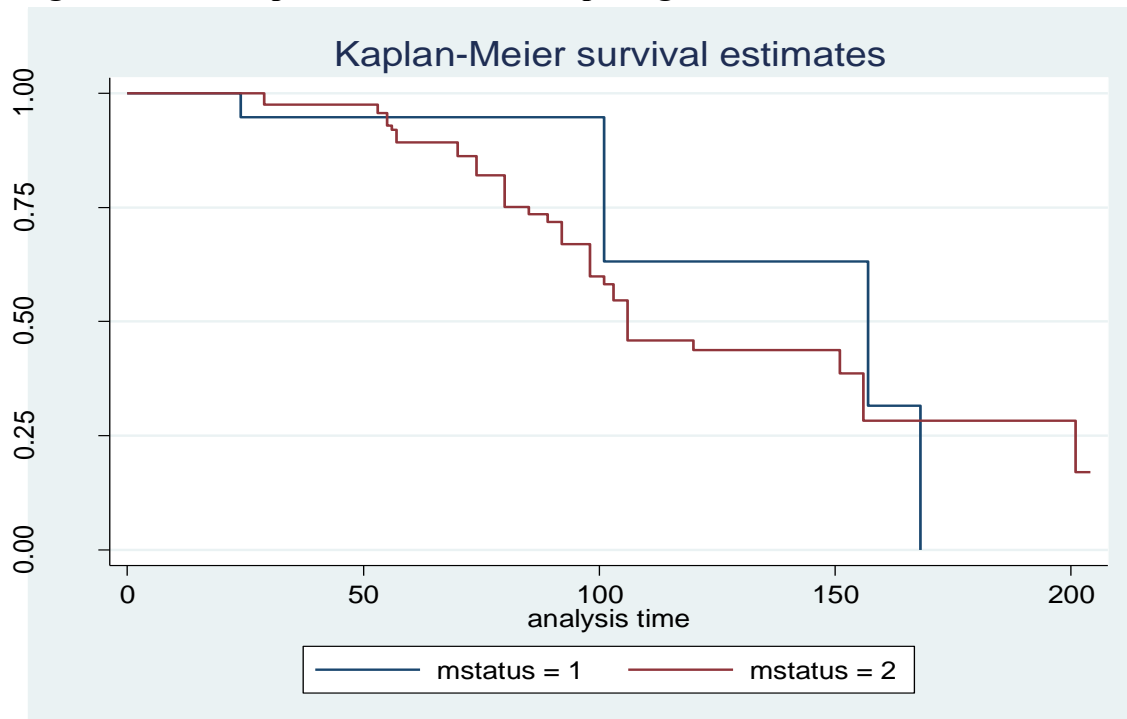


Figure 4.5 Kaplan-Meier Curve comparing Qualification Before Entry (Masters and Phd)

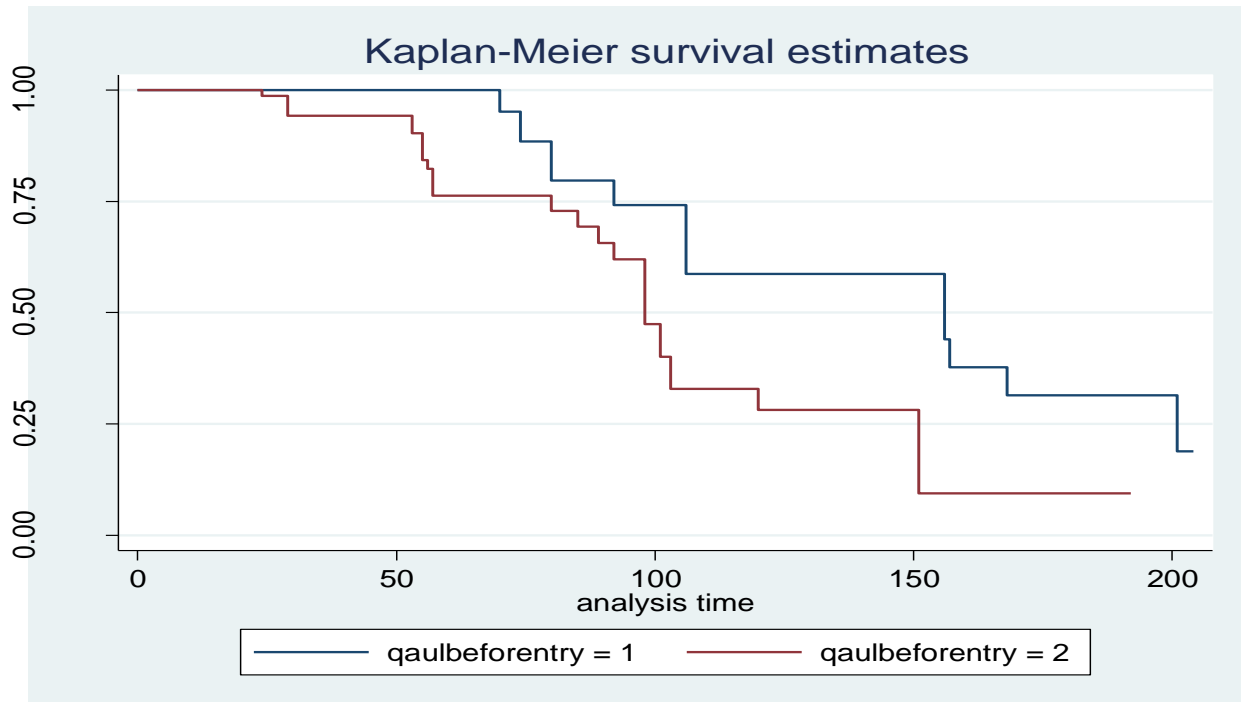


Figure 4.6 Kaplan-Meier Curve Masters Country (local and foreign)

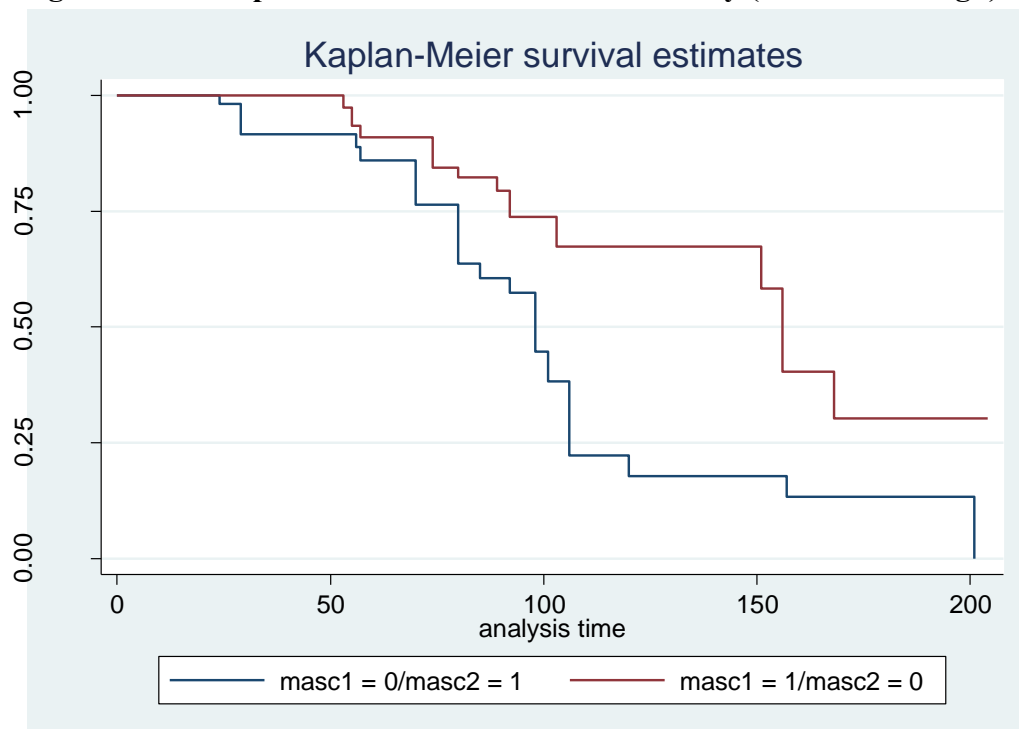


Figure 4.7 Component Bar Chart Comparing Survival Frequencies

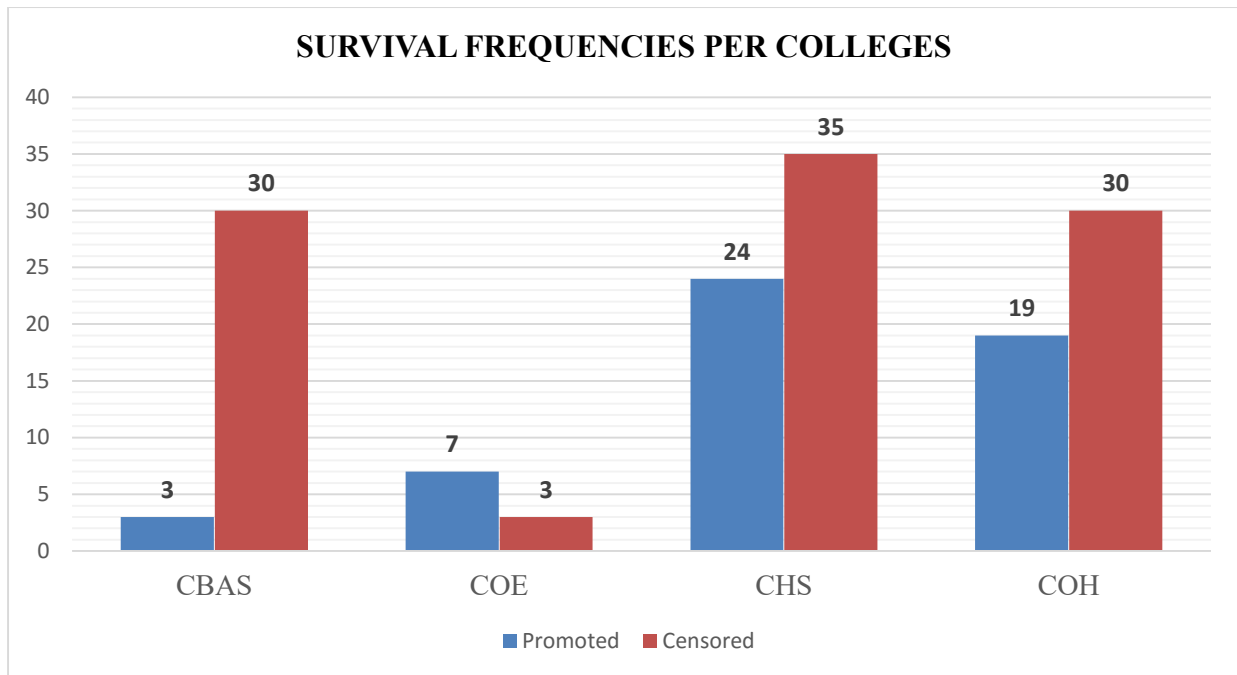


Table C1: Exponential Proportional Hazard Model

Exponential Proportional Hazard						
_t	Haz. Ratio	Std. Err.	Z	P>z	[95% Conf.	Interval]
Age	.9324922	.0465528	-1.40	0.161	.8455723	1.028347
No. of Children	1.087089	.1669344	0.54	0.587	.8045515	1.468846
No. UG committee	1.11675	.058466	2.11	0.035	1.007842	1.237427
No. Nat. committee	.9712045	.0644551	-0.44	0.660	.8527461	1.106119
No. Int. conferences	1.017101	.0199797	0.86	0.388	.9786859	1.057024
No. Books published	1.379107	.2680442	1.65	0.098	.942232	2.018543
No. Tech reports	1.058337	.0877714	0.68	0.494	.8995618	1.245137
No. Working Papers	.4822559	.1463666	-2.40	0.016	.2660324	.8742199
Male	1.03487	.4394961	0.08	0.936	.4501852	2.378922
CBAS	.4708708	.3153127	-1.12	0.261	.1267375	1.749437
COE	1.385352	.8273013	0.55	0.585	.4297755	4.465588
CHS	2.136562	.8756839	1.85	0.064	.9568524	4.770743
Local Master Deg.	.4626984	.1727993	-2.06	0.039	.2225401	.9620279
Master Cert.	.4213735	.1583583	-2.30	0.021	.2017313	.8801589
Married	.4366612	.4430542	-0.82	0.414	.0597697	3.190129
_cons	.1394444	.243477	-1.13	0.259	.0045516	4.272108

Table C2: Gompertz Proportional Hazard Model

Gompertz PH						
_t	Haz. Ratio	Std. Err.	Z	P>z	[95% Conf.	Interval]
Age	.7438294	.0497501	-4.42	0.000	.6524419	.8480176
No. of Children	1.181131	.2538302	0.77	0.439	.7751242	1.799803
No. UG committee	1.146483	.0749614	2.09	0.037	1.008586	1.303234
No. Nat. committee	1.084398	.0807749	1.09	0.277	.9370959	1.254854
No. Int. conferences	1.097895	.0282072	3.64	0.000	1.043978	1.154595
No. Books published	1.174624	.2610774	0.72	0.469	.7598128	1.815895
No. Tech reports	1.313299	.1310345	2.73	0.006	1.080028	1.596954
No. Working Papers	.1352223	.0629288	-4.30	0.000	.0543153	.3366467
Male	1.110431	.5121469	0.23	0.820	.4496818	2.742067
CBAS	.671273	.4859054	-0.55	0.582	.1624618	2.77362
COE	2.032406	1.477384	0.98	0.329	.4889451	8.448134
CHS	4.426404	2.169078	3.04	0.002	1.694086	11.56556
Local Master Deg.	.336131	.1530822	-2.39	0.017	.1376736	.8206661
Master Cert.	.0737134	.0404229	-4.76	0.000	.0251632	.2159372
Married	3.579587	4.748428	0.96	0.336	.265882	48.19223
_cons	10.48581	22.15982	1.11	0.266	.1666271	659.87
/gamma	.0408413	.0052068	7.84	0.000	.0306362	.0510463

Table C3: Exponential Accelerated Failure Time Model

Exponential AFT						
_t	Coef.	Std. Err.	Z	P>z	[95% Conf.	Interval]
Age	.0698944	.049923	1.40	0.161	-.0279528	.1677417
No. of Children	-.0835035	.1535609	- 0.54	0.587	-.3844773	.2174703
No. UG committee	-.1104231	.0523537	- 2.11	0.035	-.2130344	-.0078118
No. Nat. committee	.0292182	.0663661	0.44	0.660	-.1008571	.1592934
No. Int. conferences	-.0169566	.0196438	- 0.86	0.388	-.0554577	.0215446
No. Books published	-.3214362	.1943607	- 1.65	0.098	-.7023761	.0595037
No. Technical reports	-.0566989	.0829333	- 0.68	0.494	-.2192452	.1058475
No. Working Papers	.7292803	.3035041	2.40	0.016	.1344233	1.324137
Male	-.0342756	.4246873	- 0.08	0.936	-.8666474	.7980963
CBAS	.7531715	.6696375	1.12	0.261	-.5592938	2.065637
COE	-.3259543	.5971776	- 0.55	0.585	-1.496401	.8444923
CHS	-.7591979	.4098566	- 1.85	0.064	-1.562502	.0441062
Local Master Deg.	.7706798	.3734599	2.06	0.039	.0387118	1.502648
Master Cert.	.8642357	.3758145	2.30	0.021	.1276528	1.600819
Married	.8285978	1.014641	0.82	0.414	-1.160061	2.817257
_cons	1.970089	1.746051	1.13	0.259	-1.452107	5.392286

Table C4: Weibull Accelerated Failure Time Model

Weibull AFT						
_t	Coef.	Std. Err.	Z	P>z	[95% Conf.	Interval]
Age	.0671966	.0127524	5.27	0.000	.0422024	.0921908
No. of Children	-.0244008	.055279	-0.44	0.659	-.1327455	.083944
No. UG committee	-.0411552	.0147214	-2.80	0.005	-.0700086	-.0123018
No. Nat. committee	-.004567	.0169904	-0.27	0.788	-.0378676	.0287336
No. Int. conferences	-.0191927	.0050255	-3.82	0.000	-.0290425	-.0093428
No. Books published	-.050114	.0526154	-0.95	0.341	-.1532382	.0530102
No. Technical reports	-.0714148	.0215606	-3.31	0.001	-.1136727	-.0291569
No. Working Papers	.4296884	.0922716	4.66	0.000	.2488394	.6105374
Male	-.089743	.1177105	-0.76	0.446	-.3204513	.1409653
CBAS	.0879049	.171498	0.51	0.608	-.2482251	.4240349
COE	-.1028564	.1691952	-0.61	0.543	-.4344729	.2287602
CHS	-.4063172	.1143579	-3.55	0.000	-.6304545	-.1821798
Local Master Deg.	.2593552	.1111286	2.33	0.020	.0415471	.4771633
Master Cert.	.5894836	.1112822	5.30	0.000	.3713745	.8075927
Married	-.1687485	.3204469	-0.53	0.598	-.7968129	.459316
_cons	2.282353	.4605208	4.96	0.000	1.379749	3.184957
/ln_p	1.438857	.1133115	12.70	0.000	1.21677	1.660943
P	4.215872	.4777069		3.3762 65	5.264273	
1/p	.2371988	.0268774		.18995 98	.2961853	

Table C5: Lognormal Accelerated Failure Time Model

Lognormal AFT						
_t	Coef.	Std. Err.	Z	P>z	[95% Conf.	Interval]
Age	.0631878	.0167755	3.77	0.000	.0303085	.0960672
No. of Children	-.0325172	.0435581	-0.75	0.455	-.1178895	.0528551
No. UG committee	-.0353674	.0177815	-1.99	0.047	-.0702185	-.0005163
No. Nat. committee	-.0083514	.0257395	-0.32	0.746	-.0587999	.0420971
No. Int. conferences	-.0138757	.0063899	-2.17	0.030	-.0263997	-.0013517
No. Books published	-.0983762	.0685158	-1.44	0.151	-.2326646	.0359123
No. Technical reports	-.0455687	.0295265	-1.54	0.123	-.1034395	.0123021
No. Working Papers	.2914563	.0831859	3.50	0.000	.1284149	.4544977
Male	-.160998	.1403788	-1.15	0.251	-.4361353	.1141394
CBAS	.32779	.2126542	1.54	0.123	-.0890046	.7445845
COE	.1179823	.2364633	0.50	0.618	-.3454773	.581442
CHS	-.2211004	.114772	-1.93	0.054	-.4460495	.0038486
Local Master Deg.	.2984285	.1211802	2.46	0.014	.0609197	.5359373
Master Cert.	.4844196	.1109156	4.37	0.000	.267029	.7018103
Married	.2626636	.3834816	0.68	0.493	-.4889465	1.014274
_cons	1.827814	.6166586	2.96	0.003	.6191857	3.036443
/ln_sig	-.9718316	.0968262	-10.04	0.000	-1.161608	-.7820557
Sigma	.3783893	.036638		.3129 826	.4574646	

Table C6: Log-logistic Accelerated Failure Time Model

Log-logistic AFT						
_t	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
Age	.062503	.0139555	4.48	0.000	.0351508	.0898552
No. of Children	-.0374808	.0449732	-0.83	0.405	-.1256267	.0506652
No. UG committee	-.0384276	.0156427	-2.46	0.014	-.0690867	-.0077685
No. Nat. committee	-.008604	.0254653	-0.34	0.735	-.058515	.0413071
No. Int. conferences	-.0139797	.0052696	-2.65	0.008	-.0243079	-.0036515
No. Books published	-.0536088	.0626103	-0.86	0.392	-.1763228	.0691051
No. Technical reports	-.0569621	.025188	-2.26	0.024	-.1063297	-.0075945
No. Working Papers	.2868069	.088472	3.24	0.001	.1134049	.460209
Male	-.1801953	.1296498	-1.39	0.165	-.4343042	.0739136
CBAS	.1424115	.1837422	0.78	0.438	-.2177166	.5025396
COE	-.0255293	.1963152	-0.13	0.897	-.4103001	.3592414
CHS	-.3658238	.1223853	-2.99	0.003	-.6056945	-.125953
Local Master Deg.	.2542975	.1163737	2.19	0.029	.0262092	.4823858
Master Cert.	.5318801	.1137341	4.68	0.000	.3089655	.7547948
Married	.2240577	.5410827	0.41	0.679	-.8364449	1.28456
_cons	2.024164	.6624612	3.06	0.002	.7257642	3.322564
/ln_gam	-1.635123	.1156449	- 14.14	0.000	-1.861783	-1.408463
Gamma	.1949284	.0225425		.1553 9	.2445187	

APPENDIX B**STATA CODES**

```

use "C:\Users\JULIUS\Desktop\desktop folder\keyexcel.dta", clear
    stset time, failure(event==1) scale(1)
    tabulate gender, gen(Gn)
    tabulate college,gen(coll)
    tabulate mstatus,gen(msta)
    tabulate mascountry,gen(mastc)
    tabulate qaulbeforeentry,gen(qualbef)
    sts graph
    ltable time event, survival
    sts graph, by(gender)
    sts graph, by(Gn1 Gn2)
    sts graph, by(mastc1 mastc2)
    sts graph, by(qualbef1 qualbef2)
    sts graph, by(msta1 msta2)
    sts graph, by(Gn1 Gn2)
    sts graph, by(coll1 coll2 coll3 coll4)
    stcox age children ugcommittee natcommittee intconfatt bks techreport
workingpaper Gn1 Gn2 coll1 coll2 coll3 coll4 mastc1 mastc2 qualbef1
qualbef2 msta1 msta2
    estat ic
    streg age children ugcommittee natcommittee intconfatt bks techreport
workingpaper Gn1 Gn2 coll1 coll2 coll3 coll4 mastc1 mastc2 qualbef1
qualbef2 msta1 msta2, dist(exponential)
    estat ic
    streg age children ugcommittee natcommittee intconfatt bks techreport
workingpaper Gn1 Gn2 coll1 coll2 coll3 coll4 mastc1 mastc2 qualbef1
qualbef2 msta1 msta2, dist(weibull)
    estat ic
    streg age children ugcommittee natcommittee intconfatt bks techreport
workingpaper Gn1 Gn2 coll1 coll2 coll3 coll4 mastc1 mastc2 qualbef1
qualbef2 msta1 msta2, dist(gompertz)
    estat ic
    streg age children ugcommittee natcommittee intconfatt bks techreport
workingpaper Gn1 Gn2 coll1 coll2 coll3 coll4 mastc1 mastc2 qualbef1
qualbef2 msta1 msta2, dist(exponential) time

```

```
estat ic
streg age children ugcommittee natcommittee intconfatt bks techreport
workingpaper Gn1 Gn2 coll1 coll2 coll3 coll4 mastc1 mastc2 qualbef1
qualbef2 msta1 msta2, dist(weibull) time
estat ic
streg age children ugcommittee natcommittee intconfatt bks techreport
workingpaper Gn1 Gn2 coll1 coll2 coll3 coll4 mastc1 mastc2 qualbef1
qualbef2 msta1 msta2, dist(lognormal)
estat ic
streg age children ugcommittee natcommittee intconfatt bks techreport
workingpaper Gn1 Gn2 coll1 coll2 coll3 coll4 mastc1 mastc2 qualbef1
qualbef2 msta1 msta2, dist(loglogistic)
estat ic
```

APPENDIX C

UNIVERSITY OF GHANA

DEPARTMENT OF STATISTICS

**QUESTIONNAIRE ON DETERMINANTS OF PROMOTION OF UNIVERSITY
OF GHANA LECTURERS: A SURVIVAL ANALYSIS APPROACH**

This questionnaire is designed solely for academic purpose. Its aim is to collect data on Determinants of Promotion of University of Ghana Lecturers: A Survival Analysis approach.

Answers provided in this questionnaire will be treated with outmost confidentiality and no part will be shared with a third party. Thank you.

PLEASE TICK OR FILL IN THE BLANKS AS APPROPRIATE

1. Gender: Male

Female

2. Age: Month and Year of birth.....

3. Which religion do you practice?

Christianity

Islam (Muslim)

Traditionalist

Other (specify).....

4. What is your current marital status?

Single

Married

5. Current place of residence

On campus

Off campus

6. Number of children (if any).....

7. Children

CHILDREN	MONTH AND YEAR OF BIRTH
First child	
Second child	
Third child	
Fourth child	
Fifth child	

8. Country of Origin.....

9. Region/Province of Origin.....

10. Year of Employment.....

11. College of affiliation:

College of Basic and Applied Sciences

College of Education

College of Health Sciences

College of Humanities

12. Department.....

13. What was your qualification at time of employment?

MSC/MA/MPHIL

PHD

14. Schools attended

QUALIFICATION	UNIVERSITY	COUNTRY
Undergraduate		
Postgraduate (MSc/MA/MPhil)		
Postgraduate (PhD.)		
Post doctorate		

15. What was your level at time of employment?

Assistant Lecturer

Lecturer

Senior Lecturer

Professor

16. Have you earned promotion(s)?

Yes

No

17. Did you serve a probation period?

Yes []

No []

18. What was the duration of probation period?years

If you answered "yes" to Question 17, please answer the next question

19. Promotion(s) and dates earned

PROMOTION	DATE (DAY/MONTH/YEAR)
First Promotion	
Second Promotion	
Third Promotion	
Fourth Promotion	

20. Please answer the following as they apply to your promotion.

	1st promotion	2nd promotion	3rd promotion	4th promotion
Number of publications				
a. Books				
b. Book chapters				
c. Conference proceedings				

d. Peer-reviewed journals e. Technical reports f. Working paper g. Patent				
What was your academic qualification? e.g. MPhil, PhD, etc.				
Number of courses taught				
How many University of Ghana Statutory Committees have you served on?				
Number of national and international committees you have served on.				
Number of International Conferences attended.				

Number of International conferences in which you participated (e.g. gave a presentation or talk)				
Number of MSc. supervision				
Number of PhD supervision.				
Do you engage in any extra work aside teaching at UG?	Yes [] No []	Yes [] No []	Yes [] No []	Yes [] No []
Place of Residence prior to promotion.	Campus [] Off campus []	Campus [] Off campus []	Campus [] Off campus []	Campus [] Off campus []
Marital status prior to promotion.	Divorced [] Married [] Single [] Widowed []	Divorced [] Married [] Single [] Widowed []	Divorced [] Married [] Single [] Widowed []	Divorced [] Married [] Single [] Widowed []