

Kwame Nkrumah University of Science and Technology



**SURVIVAL ANALYSIS OF DROPOUT RATE OF BENEFICIARIES OF
COMPASSION INTERNATIONAL PROGRAMME**

By

EMMANUEL ACQUAH

PG7095616

A THESIS SUBMITTED TO THE DEPARTMENT OF MATHEMATICS IN
PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE AWARD OF
MASTER OF SCIENCE IN APPLIED STATISTICS
(MSC. APPLIED STATISTICS)

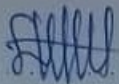
September 30, 2019

Declaration

I hereby declare that this submission is my own work towards the award of the MSc. degree and that, to the best of my knowledge, it contains no material previously published by another person nor material which had been accepted for the award of any other degree of the university, except where due acknowledgement had been made in the text.

EMMANUEL ACQUAH

(Student PG 7095616)

..........

Signature

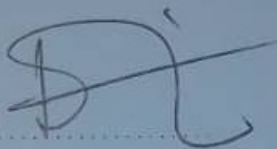
.....29/10/2019.....

Date

Certified by:

DR. S.K. APPIAH

(Supervisor)

..........

Signature

.....29/10/2019.....

Date

Certified by:

PROF. A. O. ADEBANJI

(Head of Department)

.....

Signature

.....

Date

Dedication

This work is dedicated to my loving wife Rita Esi Nyansima Acquah, my children Lordina Ewuraesi Onomabah Acquah, Emmanuella Nanayaa Kuokor Acquah and Emmanuel Acquah Jnr who with their love and care encouraged me in my educational endeavors.

KNUST



Abstract

Compassion International is Christ centered, child focused and church based organization committed to helping and supporting their beneficiaries (children) through local churches in the community. For some time now, Compassion International have been experiencing high rate of dropouts of its programme in spite of numerous benefits for the registered children. The study sought to propose a statistical model for the Compassion International Ghana beneficiaries' dropout rates using the methodology of survival analysis modelling approach. The bio-data on the beneficiaries were extracted from the database of Compassion Connect of Compassion International Ghana. The survival rates of the beneficiaries were traced at the end of each year. It was found out that the dropouts of beneficiaries were influenced by the location of beneficiaries, their gender and age group. Also, beneficiaries from rural communities were more likely to dropout from the programme and that females were more likely to dropout earlier than their male counterparts. This research will be of relevance to Compassion International Ghana and its affiliated stakeholders in planning of necessary interventions to eradicate or reduce dropout rates from such sponsored programmes.

Acknowledgements

This work will not have seen the light of day without the invaluable help of GOD and the support of my Supervisor, Dr. S. K. Appiah, who directed and guided me throughout this study. Special mention must also be made of Dr. Emmanuel Carsamer of University of Education, Winneba who provided me with the SPSS software for implementing data analysis. I acknowledge Padmore Baffour Agyapong, former Country Director of Compassion International Ghana for supporting me financially during my studies. I am most grateful to the staff of Compassion International Ghana for assisting me obtain the data for the work. To all Lecturers of Department of Mathematics and the newly created Department of Statistics and Actuarial Science, Kwame Nkrumah University of Science and Technology (KNUST) who taught, encouraged and guided me during my study, I say I am greatly indebted to you. To friends and classmates, especially Philip Teye who contributed in diverse ways to my success in this programme, I say a big thank you. My final, thanks goes to my family, for the prayers, and also bearing with me throughout this programme of study. May the good Lord richly bless you all.

Contents

Declaration	v
Dedication	v
Acknowledgment	v
List of Tables	viii
List of Figures	ix
List of Abbreviations	x
List of Notations	xi
1 Introduction	1
1.1 Background	1
1.2 Problem Statement	3
1.3 Objectives and Research Questions	4
1.4 Methodology	5
1.4.1 Data	5
1.4.2 Modelling Technique	5
1.5 Organization of study	5
2 Literature Review	6
2.1 Introduction	6
2.2 Factors of Survival of Beneficiaries	6
2.2.1 Location	6

2.2.2	Gender	7
2.2.3	Caregivers education background	8
2.2.4	Age of beneficiaries	8
2.2.5	Religious background	9
2.3	Non-Governmental Organizations (NGO's)	9
2.3.1	What is an NGO	9
2.3.2	How NGOs Help in Poverty Reduction	10
2.4	Compassion International Ghana:	11
2.4.1	Programmes in Ghana	11
2.4.2	Determinants of Beneficiaries' Dropout	13
3	SURVIVAL ANALYSIS	15
3.1	Introduction	15
3.2	Preliminary Concepts	15
3.2.1	Survival and Hazard Functions:	17
3.2.2	Censored Data	18
3.2.3	Estimation of Survival Function:	20
3.3	Cox Proportional Hazard Model	21
3.3.1	Estimation of Model Parameters	22
3.3.2	Time-Dependent Cox Model (TD-Cox)	24
4	Data Analysis and Results	26
4.1	Introduction	26
4.2	Retention (Survival) of Beneficiaries	27
4.2.1	Cluster Analysis	27
4.2.2	Log Rank Test	28
4.2.3	Gender Analysis	30
4.2.4	Survival Rate	31
4.3	Cox Hazard Proportional (Regression) Model	33
4.3.1	Cox Hazard Proportional Model for Gender	33

4.3.2	Combined Cox Hazard Proportional Model for Predictors.	36
4.4	Discussion of Results	38
5	Conclusion and Recommendations	40
5.1	Summary of Findings	40
5.2	Conclusion	42
5.3	Recommendations	42
	References	46
	Appendix A	47

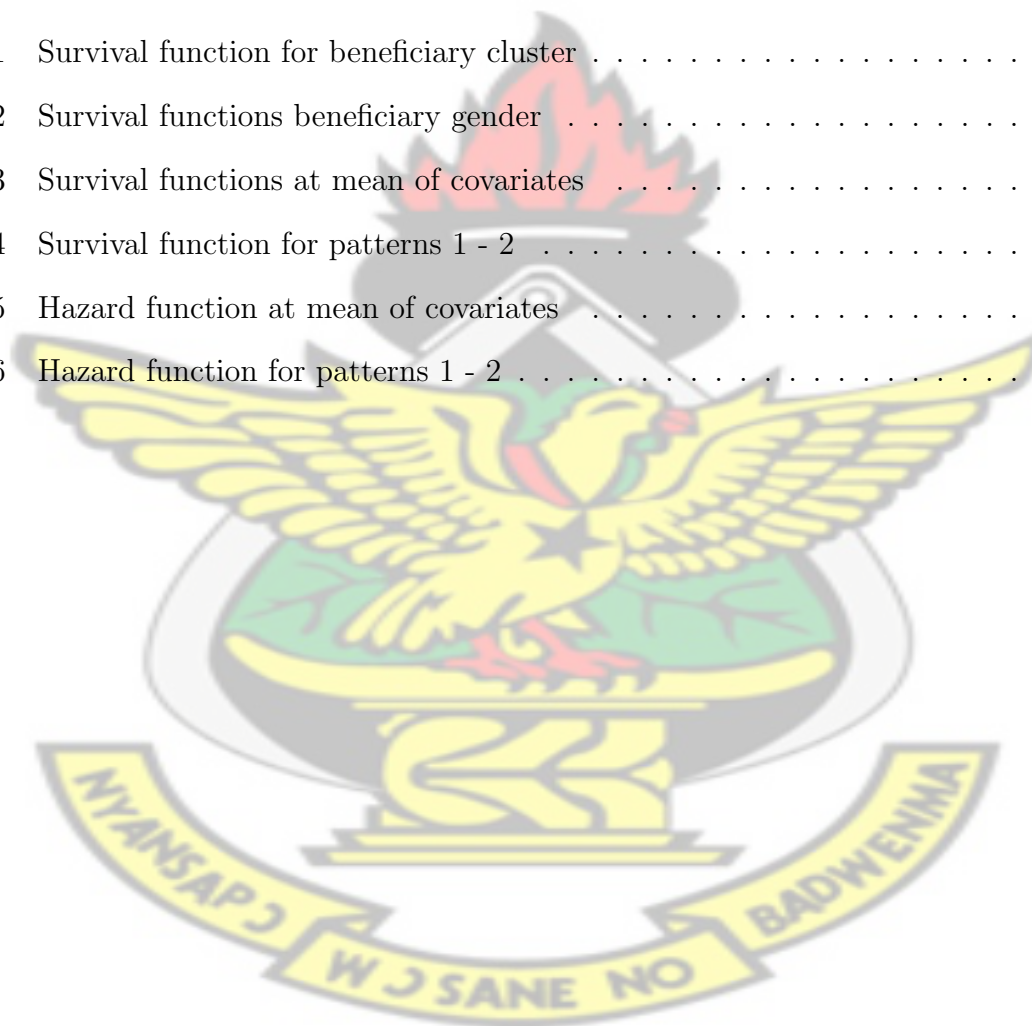


List of Tables

4.1	Categorisation of data collected	26
4.2	Means for Survival Time	28
4.3	Overall Comparisons	29
4.4	Case Processing Summary of clusters	30
4.5	Case Processing Summary of gender	31
4.6	Means for Survival Time	32
4.7	Overall Comparisons	32
4.8	Variables in the Equation for Gender	33
4.9	Cox Hazard model of dropouts predictors	37
5.1	Case Processing Summary	47
5.2	Covariate Means and Pattern Values	47
5.3	Covariate means for Location and Gender	48
5.4	Omnibus Tests of Model Coefficients	48

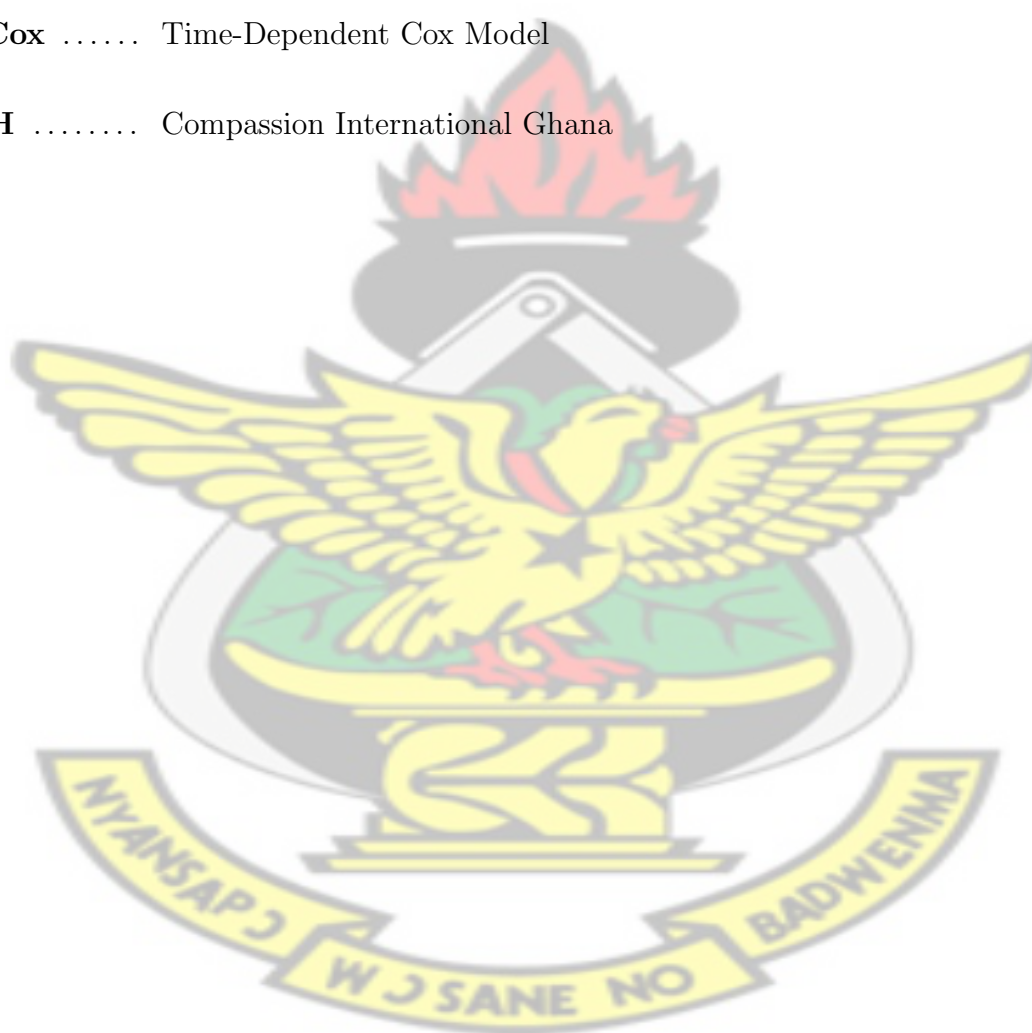
List of Figures

2.1	Compassion integrated model for Child/Youth Development.	13
3.1	Illustration of the beneficiary dropout data using survival concepts.	19
4.1	Survival function for beneficiary cluster	27
4.2	Survival functions beneficiary gender	30
4.3	Survival functions at mean of covariates	34
4.4	Survival function for patterns 1 - 2	34
4.5	Hazard function at mean of covariates	35
4.6	Hazard function for patterns 1 - 2	35



List of Abbreviations

PFM		Programme Field Manual
NGO		Non-Governmental Organization
UN		United Nations
AFT		Accelerated Failure Time
TD-Cox		Time-Dependent Cox Model
CIGH		Compassion International Ghana



List of Notations

Notation	Description
n	Number of data points
p	Number of statistics features
q	Number of time dependent features
X_i	$1 \times p$ matrix of feature vectors for subject i
$Z_i(t)$	$1 \times q$ matrix of time dependent feature vectors for subject i
T	$n \times 1$ vector of event times
C	$n \times 1$ vector of last follow up time
O	$n \times 1$ vector of observed time which is $\min(T, C)$
δ	$n \times 1$ binary vector of censored status
d_i	Number of events occurred at time t_i
$S_0(t)$	Base survival probability
$S(t X, Z(t))$	Conditional survival probability
$h_0(t)$	Base hazard rate
$h(t X, Z(t))$	Conditional hazard probability
β	$p \times 1$ vector of Cox regression coefficient
$L(\beta)$	Maximum likelihood function for β

Chapter 1

Introduction

1.1 Background

Education is a process that shapes our personality and makes us who we are (Chugh, 2004). The process of education begins even before we were born and goes on till we take our last breath. It is what makes us different from any other species on this planet. There are two ways in which people perceive education. One is where students are enrolled in an institution, get a degree, apprentice, diploma among others to get a job. On the other hand, education means learning new ways, and it refines our ability to think, decide, act, take responsibility of our actions, unlearn and relearn. Education frees us from impression of being the center of the universe (Chugh, 2004). Nelson Mandela, the former South Africa President once said, “Education is the most powerful weapon which can be used to change the world”. Education has the power to transform the world. It is not just teaching and learning, but also awakening from sleep of ignorance, illusion, misconception among others. Aristotle equated an uneducated person to a dead person and that clearly tells us how important education is for human life. It should therefore be compulsory and be the birth right of every individual. Education plays a vital role in the development of a nation and improves the standard of living of the people. In our growing world, it is perceived that a college degree enables one to obtain a satisfying job that pays well which will afford him/her a good home, food, clothing and other needs. Finally, education stimulates self-confidence and prepares us with the things we need to adequately fit in today’s world. It makes us more independent and knowledgeable of what is happening in the world (Upchurch and McCarthy, 1990). Upon the numerous benefits of education discussed above, some students still dropout from educational programmes due to the various problems including: (i) Poor parenting or

guidance: Parenting entails caring, protection, guidance and provision of basic needs for a child's up keep in order for him or her to be properly equipped to meet the challenges of life (Odia and Omofonmwan, 2007) . In desperation, many parents have decided to bring in additional innovation by way of not only involving in encouraging, but also finance activities in and around examination venues to perpetrate examination malpractice in order to enhance the chances of their wards passing examination and qualifying to higher institutions. Interview with two sets of university students from various departments revealed that their parents' influence over their choice of course of study has negative effect on their level of performance. Most of the uneducated parents do not have time for their children. They do not even see the reason why they should waste money on their children at school. (ii) Another problem of dropout is poverty. According to the World Development Report 2001, 22 percent of the children in the United States live at or below poverty level. Students living at or below poverty level are more likely to have the highest dropout rates. Students who do not get enough food or sleep tend to perform lesser than their full academic potentials. Due to the challenges associated with education that lead to students' dropout from school, the government of Ghana has put some policies in place to mitigate these challenges. One of the policies is the School Feeding Programme. It is defined by the World Bank as a "targeted social safety net that provides both health and educational benefits to the most vulnerable children, thereby increasing enrolment rates, reducing absenteeism, and improving food security at the household level" (Government, 2011). Beyond improvements in access to food, the school feeding programme has also had a positive impact on gender equity, educational status and nutritional status. Beneficiaries of this programme are children in low-income mostly in deprived communities. The programme is being funded by international organizations such as the United Nations World Food Programme, the World Bank or McGovern-Dole International Food for Education and Child Nutrition programme, (Government of Ghana 2011). Apart from school feeding, governments, NGOs and philanthropist have provided financial aid to support students, girl child among others. All these interventions are put in place to mitigate school dropouts and absenteeism.

1.2 Problem Statement

Governments all over the world aim to improve educational and also increase enrollment, retention and high achievement. Students' retention is increasingly becoming a problem to all and sundry. The negative impacts of student dropout are clear to parents and institutions of learning. The positive effect of decreasing student dropout is also self-evident including higher tendency of having a better career and higher standard of living for students. Not only from a student perspective but also college rankings, federal funding agencies and state appropriation committees are all directly influenced by student retention rates (Chugh, 2004). Thus, the higher the student retention rate, the more likely the educational institution is positioned higher in the ranking, secures more government funds, and has easier path to programme accreditations. In view of these reasons, Directors in higher education want to outline and implement strategies to increase students' retention. The objective of the study is to provide a detailed analysis of the students' retention problem and predict the risk of dropout from educational institutions by beneficiaries of Compassion International. Methods that are currently being used for the problem of student dropout are the standard preliminary statistical approaches. The study takes advantage of multiple kinds of information about different aspects of student's characteristics and efficiently utilize them to make a personalized decision about the risk of dropout for a particular student. In this regard, we propose survival analysis modelling approach which addresses "time to event occurrence". In such cases, the reliable estimation of risk at early stage of student education is very important. The basic idea here is to utilize the survival analysis method at early stage of Compassion International Ghana Programme to predict beneficiary's success (Ameri, 2015). Compassion International Ghana is one of the numerous Non-Government Organizations (NGOs) in the world which aims at poverty reduction among the rural poor. Compassion International Ghana found its way to Ghana in 2005. Since then, it has partnered with over 200 churches and communities in the Central, Greater Accra, Eastern, Volta, Ashanti, Brong-Ahafo and Western Regions to help reduce poverty by transforming the lives of children through

their holistic child development programme. The holistic child development programme spans over four key areas, namely:

- Cognitive development of the beneficiary, which aims at helping him/her to become economically self-supporting. This is to make sure the beneficiary is able to complete at least Senior High School and acquire an income generating skill.
- Spiritual development of the beneficiary: As a Christian organization, it makes sure that the beneficiary is able to exhibit commitment to the Lordship of Jesus Christ. By so doing, the beneficiary is helped to know and understand the Bible, know how to pray and support church activities, and also know the importance of the death of Jesus Christ.
- Socio emotional development of the beneficiary: For the beneficiary to integrate well in the society, Compassion International Ghana helps them through this programme model to be able to exercise self-management, interact with others in a healthy and compassionate manner, and to apply self and social awareness in making responsible life choices.
- Physical development of the beneficiary: The programme aims at providing 80% preventive and 20% curative measures against diseases a registered beneficiary may go through. They make sure that the beneficiary understands the need to care for his or her physical body, chooses good nutrition and takes responsible life choices about health and sexuality (Compassion, 2017).

Hundreds of beneficiaries are registered to benefit the Compassion International programme at various centers every week but upon all the benefits from the programme, some beneficiaries' dropout from the programme.

1.3 Objectives and Research Questions

The purpose of the study is:

1. To examine the determinants of beneficiary dropouts (exit) from the Compassion International Ghana programme.
2. To examine the effect of supervision, counselling, and time of funds release on dropout rates.

1.4 Methodology

1.4.1 Data

The bio data collected on the beneficiaries were obtained from Compassion International Ghana database. They include: age, gender, location, caregiver occupation and religion.

1.4.2 Modelling Technique

The study proposed a Kaplan-Meier estimate and Cox-Hazard regression model to examine the time it takes a dropout of a beneficiary to occur, also establish the relationship between survival and one or more predictors, usually termed as covariates (Kaplan and Meier, 1958). Generally, the study estimated the survival function, hazard function and Cox proportional hazard.

1.5 Organization of study

The study consists of five chapters. Chapter one is the introduction which includes background of the study, the problem statement, the objectives, and research methodology. Review of relevant literature is presented in Chapter two. Chapter three gives the detailed outline of survival analysis, the methodology of the study. Chapter four presents the data analysis and discussion of the results. Chapter five concludes the study with summary of the findings of the analysis and recommendations.

Chapter 2

Literature Review

2.1 Introduction

This chapter reviews the relevant literature for the work. School dropout is a result of many factors that come into play and it also has effects on the society as a whole. According to (Milligan et al., 2004), poverty is positively associated with school dropout. Moreover, economic growth is affected by high drop out rates according to (Hanushek and Wößmann, 2007).

2.2 Factors of Survival of Beneficiaries

2.2.1 Location

The location of beneficiaries significantly affects their survival on sponsorship programmes. Those from disadvantaged backgrounds are less likely to survive in the programmes. According to Saeed et al., (2016), 32.3% of the total respondents claimed they were unable to survive in programme due to their disadvantaged backgrounds. It is clear to point out that, parents of students especially in deprived areas spent more of their incomes on funerals, social gathering and expensive clothes and shoes while they neglect spending or having time on their children's welfare which ultimately affect their wards survival in programme. For example, if a child is in programme and he/she must attend a meeting, the caregiver will attend to social gathering issue instead of making sure the child should attend. In some of these deprived areas too, beneficiaries' survival on programmes are hindered by their desire to gain 'quick money'. These children would prefer engaging in economic activities such as farming and small-scale mining to earn a living than to spend so many years in sponsorship programme. Furthermore, Okumu et al. (2008), found that,

the survival of students diminishes as they move from urban areas to the rural areas. This can be explained by the access to facilities in the urban areas than in the rural areas coupled with the high motivation to stay on a programme in urban areas than in rural areas. Location can also be explained from the view of the location of beneficiaries' houses to where the programme center is. It has been found that, students who are residents near the programme center areas are less likely to survive as compare to those who must travel long distances to the center (Finn et al., 1988).

2.2.2 Gender

Several findings have been made in line with the survival rate between male and female children in the programme. Saeed, et al., (2016), observed that in the 2012/2013 academic year, 391 males out of a total of 12,086 representing 3.24% dropped out of school while as much as 789 females out of 10,449 (7.55%) dropped out of school. The implication of this is that, a number of the students (1180) could not survive in the educational program and subsequently had to drop out of school. This is in line with the findings of research by Holmes (2003) and Adamson (1996). It was observed that, female beneficiaries had less survival rate on sponsorship programmes than male counterpart. Female beneficiaries' survival on sponsorship programmes may be less especially in rural areas. The reasons could be attributed to teenage pregnancies and early marriage. However, the survival of male students in the 2013/2014 academic year increased while that of female students decreased. The enrolment of students was 11,346 and 9,792 for male and female students respectively. The study observed that, 7.08% of students dropped out of school which shows that students were not able to survive in the educational program in the 2012/2013 academic year. Male students' survival may be affected by preference of work to school. Male students, especially those from disadvantaged backgrounds would prefer to work to support their parents than to spend many years in school. One other reason may also be male students' likelihood to engage themselves in delinquent attitudes such as runaway, aggressiveness, drug use and sale (Hartnagel and Krahn, 1989).

2.2.3 Caregivers education background

Parents' educational background affects beneficiaries' survival in programmes. This includes both fathers and mothers educational background. It has been found that, children whose parents received more education were likely to be enrolled and survive on sponsorship programmes. Choudhury (2006) observed a significant direct relationship between father's educational level and students' survival in programmes. The study concludes that a step above the level of a father's education, leads to an increase in his ward's survival in educational programmes like Compassion International Ghana's sponsorship programmes. (that is, 16% reduction in dropout). The reason is that, he can foresee benefits the child can derive from the programmes. On the other hand, it was observed that there was little significant effect of mother's educational level on their ward's survival on educational programmes. Holmes (2003) observed an interesting finding about parents' educational level on the survival of their children in educational programmes. He noted that, father's educational level goes a long way to enhance a boy's survival in school and mothers educational level enhances that of girls. This implies that male students draw inspiration mostly from their fathers and girls from their mothers.

2.2.4 Age of beneficiaries

Lazar et al. (1982) studied the effects of pre-school educational programmes on students' survival. Their research was mostly targeted at children from disadvantaged backgrounds. They found that, educational programmes at the pre-school that were mostly cognitively oriented affected students to perform better at their later schooling. They added that, cognitive oriented pre-school educational programmes enabled students to be more likely to be employed and also such students were not held back while their peers moved forward or were not assigned special education. The research made it clear how educational programmes at the pre-school affected students' survival. The study compared such students with those in the controlled group (i.e. children without pre-school). The study asked these children to comment on something they have done that they felt proud of. Distinc-

tively, the study observed that pre-school graduates were more likely to come out with answers that indicated what they have achieved. Furthermore, mothers of pre-school graduates reported that they were more satisfied with their children in terms of their school work and their employment aspirations. Okumu et al. (2008), observed that, students are likely to survive in educational programmes as the age of their household heads increase. This is undoubtedly true in the sense that, as household heads grow, they appreciate the role and the merits of education and as a result, they tend to influence the survival of their children in school through pieces of advice.

2.2.5 Religious background

The religious background of beneficiaries has been found to affect beneficiaries' survival in educational programmes. According to Choudhury (2006), there was a 1.9 times likelihood that children who are Muslims could not survive on Christian programmes as compared to their colleagues who are protestant. The reasons are obvious, the Muslims think that, the Christians will convert their children from their Islamic believe.

2.3 Non-Governmental Organizations (NGO's)

2.3.1 What is an NGO

The non-governmental organization (NGO) came into existence in 1945 because of the need for the United Nations (UN) to distinguish in its Charter between participation rights for intergovernmental specialized agencies and those for global private organizations. At the UN, all types of private organizations can be recognized as NGOs. The major factor being their independence from government control, and not seeking to challenge governments either as a political faction or but a slight focus on human rights, non-profit-making and non criminal (Lawry, 2009). The composition of NGOs varies extensively based on their focus. They can be world-wide hierarchies, with either a comparatively powerful principal authority or a looser centralised layout. On the other hand, they may be based in a single country and operate transnational (Lewis and Kanji, 2009). Likewise, NGOs

are created to address specific needs in the society. While some are religion oriented, others are also regionally based, technically specialized, or community-based organizations. The global community of NGOs encompasses a wide range of organizations, which can usually be categorized into four basic types:

- Humanitarian aid, both emergency response and developmental
- Advocacy groups such as Human Rights Watch and Amnesty (International),
- Faith-based groups are humanitarian aid groups originally founded on faith principles but do not use religion as a basis of giving aid. For example, Compassion International is a Christians NGO but their child sponsorship covers Christians and non-Christians.
- Missionary aid groups seek to “spread the word” using humanitarian aid to access to those who need conversion, more or less conditional aid (Lawry, 2009).

2.3.2 How NGOs Help in Poverty Reduction

NGOs approach towards poverty reduction is different. Some of the approaches are: community based programme, church based programme, agriculture based programme and partnership with government agencies. Most of these NGOs target children and women as most effective group in the society that can help reduce poverty. This is because women paid and unpaid work is crucial for the survival of poor households. They produce and process food for the family and are the primary caretakers of children, elderly and the sick. Their income and labor are mainly directed towards children’s education, health and well-being (Gupta and Singh, 2016). In addition, in poverty, women and children suffer more than men (Khan, 2001). In this section, we will look at the approach of some NGOs which work hand in hand with communities, families, partners and supporters to make sure that children enjoy good health, are well educated, experience the love of God and their neighbors are cared for and protected. For instance, World Vision International helps in terms of disaster management and advocacy for the poor in the community especially the children. They believe that if a community is transformed, a child will

be transformed, and poverty will reduce, (Tetteh, 2016). Second, Plan International works in similar manner as World Vision International, partner with communities to transform lives of children with the aim of reducing poverty. Their work is connected to The United Nations Convention on Rights of the Child which spells out the human rights of children, including the right to survive, develop to the fullest, be protected from harmful influences, abuse and exploitation, and participate fully in family cultural and social life. The Hunger Project approach to poverty reduction stands on three essential pillars; empowering women as key change agents, mobilization of entire community into self-reliant action and to foster effective alliances to engage local authorities.

Furthermore, the Hunger project strategy mobilizes cluster of rural villages into epicenters which band together 5000-15000 people to carry out community- led integrated strategies to meet basic needs. Women and men in 121 epicenters create and run their own development programme (Von Grebmer et al., 2016). Finally, the Compassion International partners with local churches in releasing children from poverty.

2.4 Compassion International Ghana:

2.4.1 Programmes in Ghana

Compassion International is a Christ centered, child focused, and church-based organization committed to the local churches. They have chosen to partner with the local churches because it believes God has instituted the church and placed it in the communities to minister. It also believes that by transforming one child or releasing a child from poverty, it will have rippling effect on the child's family and thereby the entire community, (Compassion, 2017). They partner with the local churches in the community and assist the local churches to support the beneficiaries through a holistic child development programme. The holistic child development method encompasses cognitive, spiritual, socio emotional and physical development of all registered beneficiaries (children). Sometimes the Compassion's strategy also goes a long way to addressing the larger context of the child which includes the child's immediate family and the community he resides, (Compassion, 2017).

Compassion International Ghana support partner churches financially and in capacity building. Compassion Ghana started its operations in the Greater Accra Region in July, 2005 with 7 Implementing Church Partners (ICPs) and 1000 beneficiaries. Currently, it is operating in 7 out of the 16 regions in Ghana: The Greater Accra, Ashanti, Central, Western, Brong Ahafo, Eastern and Volta Regions, (Compassion International Ghana, 2017 annual report). Presently, Compassion International Ghana (CIGH) is administering the Child Development through Sponsorship Programme (CDSP). This is done through home and center based programmes. Children between the ages of 3 and 6 years are registered into this programme. Implementing Church Partner (ICP) meet beneficiaries between 4 to 6 hours every week to take them through well-structured curriculum. Meanwhile, the Child Survival Programme (CSP) beneficiaries are registered at age zero up to age one. They register the beneficiaries when the mother is pregnant. The purpose of this intervention is to reduce mortality among mothers, newborns, and infants across the globe, providing essential interventions to help children reach their first birthday. Home based implementers are trained to reach out to pregnant women, vulnerable children and their primary caregivers in the community through home visit and group activities. They save children's lives and set them on a path towards healthy development. Intervention's strategies include ensuring pregnant mothers and babies have access to prenatal care, skilled birth attendants, medical treatment, growth monitoring, immunizations and early childhood development education, (Compassion, 2017). When the beneficiary is three years he/she is transitioned into CDSP. The programme ensures all the beneficiaries benefit from improved learning opportunities (school fees, school supplies, extra tuition where necessary), access to healthcare and health education, Supplementary food (where necessary), one-on-one interaction with caring christian adults and a chance to hear about the love of Jesus Christ. These programmes are run on a model called Holistic Child and youth development programme model. The model consists of survival intervention, early childhood, childhood and the youth. Figure 2.1 shows the Compassion integrated model for Child / Youth Development programme.

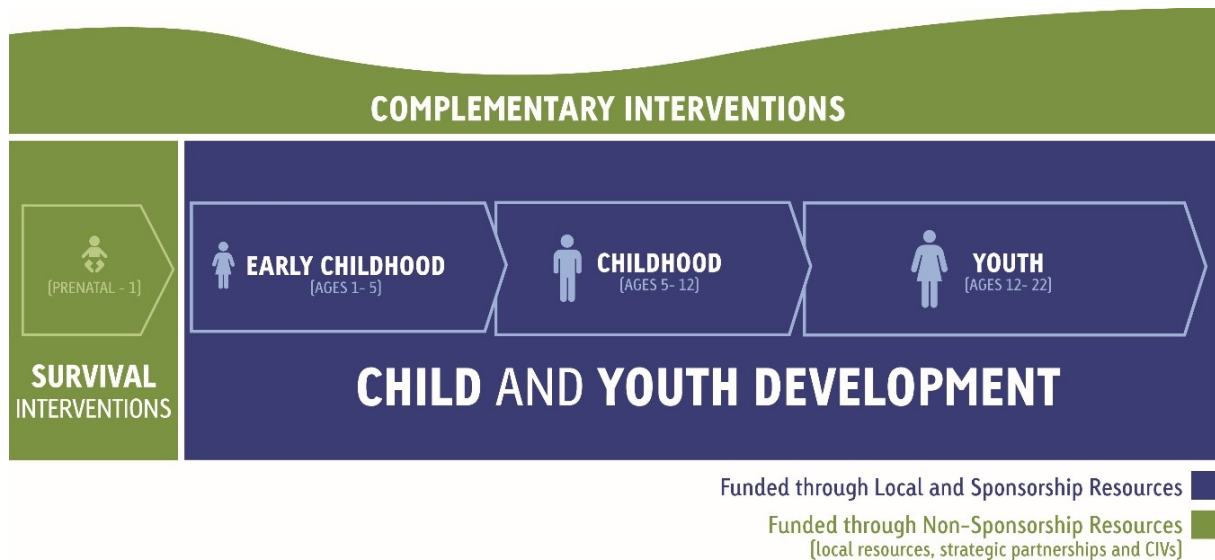


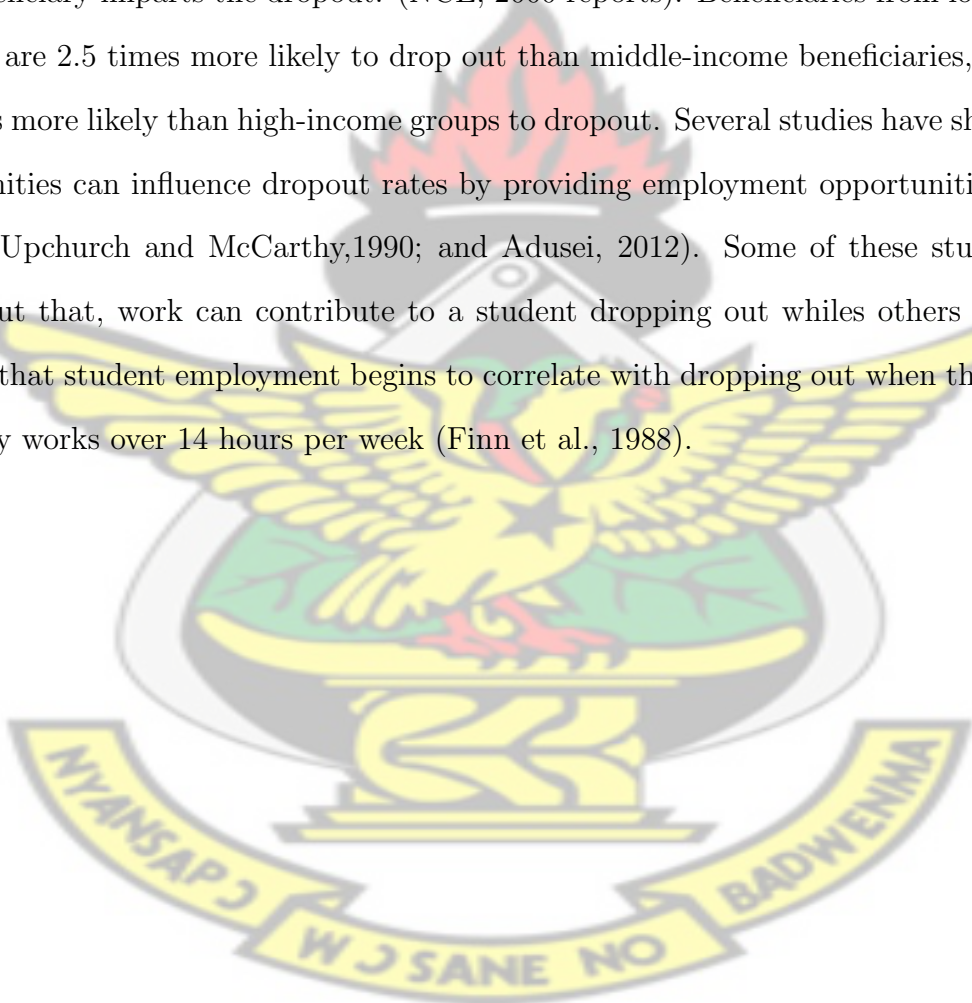
Figure 2.1: Compassion integrated model for Child/Youth Development.

Compassion (2017)

2.4.2 Determinants of Beneficiaries' Dropout

The determinants of the beneficiaries drop out are the relocation, interest and socio-economic. Relocation is one of the major causes of beneficiary dropout from the compassion programme. This happens when caregivers move to another area where the compassion international related programme is not available. The purpose of CIGH intervention is to serve the children within the local community. That is, the distance from a beneficiary's house to the project center should not be more than three kilometers. Compassion International Ghana, 2017 annual report revealed that 50% of the dropouts within the year was as a result of relocation, (Compassion, 2017). Most beneficiaries drop from the programme because of loss of interest in the programme. CIGH programme is designed in such a way that, the implementing church partner (ICP) must contextualize its own activities to suite their environment. So when the ICP do not involve the beneficiaries in planning, they end up planning something which does not interest the children. The project level factors also play a role in increasing pressures to dropout such as teacher absenteeism, project location and poor quality programming, (Hartnagel and Krahn, 1989). Availability of educational infrastructure are related to quality in terms of human re-

sources. Teaching approaches and behavior of some teachers influence beneficiaries to drop out, Smith, (2003) found out that in some schools in Zimbabwe's Southern Province teachers did not prepare lesson, had no scheme of work and left pupils assignment, Miles (2009) argued that certain teachers still employ authoritarian types of teaching which could result in learner absenting themselves from school. Such environments in compassion programme will lead to beneficiary's loss of interest in the programme. One unchanging factor when it come to the dropout rate is socioeconomic background. A study from National Center for Education Statistics showed that, the socioeconomic status of the beneficiary imparts the dropout. (NCE, 2000 reports). Beneficiaries from low-income families are 2.5 times more likely to drop out than middle-income beneficiaries, and over 10 times more likely than high-income groups to dropout. Several studies have shown that communities can influence dropout rates by providing employment opportunities during school (Upchurch and McCarthy,1990; and Adusei, 2012). Some of these studies have found out that, work can contribute to a student dropping out while others have also showed that student employment begins to correlate with dropping out when the student regularly works over 14 hours per week (Finn et al., 1988).



Chapter 3

SURVIVAL ANALYSIS

3.1 Introduction

Survival models are used to estimate time to events of interest. The ability to model the dynamic nature of incidents is a powerful tool because in many cases it answers the question of “when” which is as important as “who”. Furthermore, it is important to identify the characteristics that led to the occurrence of events. This is very important in the beneficiary dropouts problem of Compassion International Ghana as a dropout is not an instant incident, but rather a process that totally depends on time. Hence, survival analysis procedures would be appropriate statistical tool for modelling such kind of problem (Bland and Altman, 2004). The chapter seeks to explain existing survival analysis methods and utilize them to examine why the beneficiaries of the sponsored programme of Compassion International Ghana dropout. Survival analysis methods have been developed in the field of Statistics to study the time that an event will occur and focuses on using the techniques for predicting occurrence of event (Bland and Altman, 2004). In this case multivariate analysis which explores the use of time-dependent covariates is used to examine factors that account for dropout in this case study.

3.2 Preliminary Concepts

Survival analysis is a set of techniques that can be used to analyze data where the resulting variable is the time until the occurrence of an event of interest (Cornell, 2010). This kind of set of data has three main characteristics, namely:

- The dependent variable (or response) which is the time until an event occurs;
- Censored time for observations, which means that for some subjects the event of

interest has not occurred (or not recorded) at the time the data is analyzed; and

- Predictor or attribute variables that have effect on the time to event.

One of the important characteristics of longitudinal data is that, it can be incomplete due to the challenge of follow-up on the subject continuously, which is known as censoring. This incompleteness in events or records in longitudinal data is in many ways differs from missing observations associated with routine data mining problems. Not all modelling procedures are able to appropriately tackle them. Thus, it is difficult for standard machine learning methods to model data which contains censoring. Ignoring the censored data on one hand yields suboptimal biased models because it will mean overlooking available information while on the other hand, treating censoring time as the actual event time causes the model to be inadequate. Also, unlike machine learning and data mining techniques which normally come up with single outcome prediction, survival analysis estimates the survival (failure) as a function of time (Hosmer Jr and Lemeshow, 1999). In survival analysis, subjects are usually followed over a stipulated time period and the focus is on the time at which the event of interest occurs. For the beneficiaries' dropout problem, a question that arises is that why linear regression cannot be used to model the yearly dropout. One important reason is that linear regression cannot handle the censored observations in an efficient way. Unlike regression models, survival models incorporate information from both uncensored and censored data to evaluate important features. One of the paramount aspects of survival analysis is covariate data are collected longitudinally. It appears to be regular and proper to use the covariate information that vary over time in an appropriate statistical model. This aspect of time-dependent covariates makes survival data very unique in a way that other standard machine learning methods could not handle. Despite the numerous applications of survival analysis methods in other fields such as health care, there is only a limited attempt of using these methods for the beneficiaries' dropout problem.

The remaining subsections present some basic definitions, of survival analysis after which the proposed survival analysis framework for the beneficiaries' dropout prediction is presented.

3.2.1 Survival and Hazard Functions:

Survival analysis consists of two main components. First, is the event time and the second, is the status of the event which has the occurrence information for the event of interest. With event time, we can fit it into two functions that are dependent on time, namely the survival and the hazard functions. These two functions are the critical concepts in survival analysis used to describe the distribution for times of events. For every specific time, the survival function gives the survival probability until that time. The hazard function gives the possibility that the event will occur per time unit. For instance, if we let T denote the survival time of an individual, which has density function $f(t)$, then the distribution function is given by:

$$F(t) = \int_0^t f(u)du \quad (3.1)$$

do not provide particular information about the chance of survival at a given time point. Instead, the survival, hazard, and cumulative hazard functions, which are functions of the density and distribution functions, are used (Sattar Ameri, 2015).

The survival function is defined as the probability that the event of interest has not occurred by t . The survival function can be expressed in terms of probability distribution and probability density functions:

$$S(t) = Pr(T > t) = \int_t^\infty f(u)du = 1 - F(t) \quad (3.2)$$

The hazard function is an alternative characterization of the distribution of T is given by the hazard function, or instantaneous rate of occurrence of the event, defined as

$$h(t) = \lim_{dt \rightarrow 0} \frac{Pr(t \leq T < T + dt | T \geq t)}{dt} \quad (3.3)$$

In other words, $h(t)$ is defined as the rate of the event at time t which depends on survival until time t . The numerator of this expression is given as the conditional probability that the event will occur in the interval $(t; t+dt)$ given that it has not occurred

until time t , and the denominator is the length of the interval. Dividing, we obtain a rate of event occurrence per unit time. Taking the limit as the width of the interval goes to zero, we obtain an instantaneous rate of occurrence. The conditional probability in the numerator may be expressed as the ratio of the joint probability that T is in the interval $t; t + dt$ and $T \geq t$ to the probability of the condition $T \geq t$. The former may be written as $f(t) dt$ for small dt , while the latter is $S(t)$ by definition. Dividing by dt and taking the limit yields the result;

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

In other words, the rate of occurrence of the event at duration t is the same as the density of events at t , divided by the probability of surviving through the given time without experiencing the event. $S(t)$ can also be expressed as (Ameri, 2015)

$$S(t) = e^{(-\int_0^t h(x)dx)} \quad (3.5)$$

The cumulative hazard function $H(t)$ as defined in (3.6) is the sum of the risks that an individual face going from duration 0 to t . These results show that the survival and hazard functions provide alternative but equivalent characterizations of the distribution of T .

$$H(t) = \int_0^t h(x)dx \quad (3.6)$$

3.2.2 Censored Data

One of the main features of survival data which distinguishes it from all other kinds of data is that it is often incomplete. This means that the event information for some observations is not complete and such instances are considered to be censored. There are three main different types of censoring; right, left and interval censoring. Most of censoring is right which means we may not observe the time point; thus, for some reason independent of its survival time, the individual chooses to leave the programme. In the beneficiaries' dropouts problem the event of interest is "dropout", which is considered as right censoring. Let us suppose that T_i is the survival time, we observe instead $Y_i = C_i$

(where C_i is the censoring time) since T_i may not be observed. We do know whether the data has been censored, and together with Y_i we observe the indicator variable.

$$\delta_i = \begin{cases} 1 & T_i \leq C_i \\ 0 & T_i > C_i \end{cases} \quad (3.7)$$

where left censored = event start before observation began, right censored = event ended after the end of the observation and interval censoring = event occurred within the duration of the study. So, if for individual i , $\delta_i = 0$ it is censored and if $\delta_i = 1$ it is not censored. Figure 3.1 illustrates an example of beneficiaries' dropouts problem using survival analysis. A beneficiary having shaded box indicate the year of dropout. For instance, beneficiaries with ID GH020500006, GH020500032 and GH020500057 dropped out (uncensored) at 8th, 5th and 7th year respectively and non-shaded boxes implies surviving in the programme to the end (throughout the ten years).

Beneficiary ID	Year of Dropout									
	1	2	3	4	5	6	7	8	9	10
GH020500004										
GH020500006										
GH020500030										
GH020500032										
GH020500033										
GH020500053										
GH020500056										
GH020500057										
GH020500058										
GH020500179										

Figure 3.1: Illustration of the beneficiary dropout data using survival concepts.

3.2.3 Estimation of Survival Function:

In order to estimate parameters or other kinds of estimates from survival models they can be observed as ordinary regression models in which the dependent variable is time. However, the likelihood function in the presence of censored data is more complicated. The likelihood function for a survival model is a product of the likelihood probabilities and densities, depending on whether the observations were censored or not. By definition, the likelihood function is the product of the likelihood of individual observations. It is suitable to divide the data into four categories: uncensored, left censored, right censored and interval censored denoted by “unc.”, “l.c.”, “r.c” and “i.c”, respectively. The likelihood function can be formulated as given in equation 3.8:

$$L(\theta) = \prod_{T_i \in \text{unc}} Pr(T = T_i|\theta) \prod_{i \in \text{l.c}} Pr(T < T_i|\theta) \prod_{i \in \text{r.c}} Pr(T > T_i|\theta) \prod_{i \in \text{i.c}} Pr(T_i, l < T < T_i, r|\theta) \quad (3.8)$$

where r is right censored, l is left censored and T_i is the event time

In this thesis, we only consider data with event occurring with right censored data. So the likelihood function equation (3.8) can then be written using only uncensored data as;

$$Pr(T = T_i|\theta) = f(T_i|\theta) \quad (3.9)$$

$$Pr(T > T_i|\theta) = 1 - F(T_i|\theta) = S(T_i|\theta) \quad (\text{for right censored}) \quad (3.10)$$

The analysis of survival data can be done in multiple ways. One of the common methods is the non-parametric approach which has no assumption about the form of the survival distribution. One of the well-known non-parametric estimator of survival function is the use of Kaplan-Meier method which is widely used to estimate and graph survival probabilities as a function of time. It can be used to obtain univariate descriptive statistics for survival data including the median survival time, and to match the form of the survival between two or more groups. This Kaplan Meier estimator is defined as:

$$\hat{S}(t) = \prod_{i:t(i)<t} \left(1 - \frac{d_i}{n_i}\right) \quad (3.11)$$

where d_i is the number of failures before time t , and n_i denotes the number of individuals who have neither experienced the event of interest nor have been censored, by time t . On the other hand, parametric methods assume that the underlying distribution of the survival times follow a certain known probability distribution, some of which are; lognormal, Weibull and exponential distributions. The description of the distribution of the survival times and the change in their distribution as a function of predictors is of interest. The model parameters are usually estimated using an appropriate adjustment of the maximum likelihood function (Ameri, 2015).

3.3 Cox Proportional Hazard Model

In the previous section, we discussed two categories of techniques for modelling survival data. The non-parametric method only considers time-to-event data and does not take care of any explanatory variables that might relate to the event occurrence. On the other hand, parametric (or semi-parametric) models are able to consider covariates in the analysis, but these models usually demand knowledge of the data distribution. One advantage of models like semi-parametric model which can model survival data using covariates is that, little assumption is made about time-to-event data. A widely used method in this category is Cox proportional hazard model which makes fewer assumptions than typical non-parametric methods. In particular, Cox proportional hazard model, unlike parametric models, has no assumptions about the form of the baseline hazard function although the covariates should be dependent on time.

Let T_i denote the observed time which can be either event time or censoring time for subject i . Also, let δ_i be the event status indicator; if $\delta_i = 1$, then the event has occurred and if $\delta_i = 0$, then the subject is censored. The hazard function for the Cox proportional

hazard model has the functional form

$$h(t|X) = h_0(t)e^{(\beta_1 X_1 + \dots + \beta_p X_p)} = h_0(t)e^{(\beta, X)} \quad (3.12)$$

where $h_0(t)e^{(\beta(t))}$ is the baseline hazard function at time t and $e^{\beta_1 X_1 + \dots + \beta_p X_p}$ is the risk associated with the covariates. If we take the ratio of the hazards, the baseline hazard cancels out and the hazards are proportional at any given time t , yielding the proportional hazards model. This expression gives the hazard at time t for an individual with covariate vector $X = x_1, x_2, \dots, x_p$. Therefore, the survival probability function for Cox proportional model can be formulated as

$$S(t|X) = S_0(t)e^{(\beta X)} \quad (3.13)$$

where

$$S_0(t) = e^{-\int_0^t h_0(x)dx}$$

3.3.1 Estimation of Model Parameters

Parameter estimation of the Cox Proportional hazard (or regression) model is done by maximizing the partial likelihood function as opposed to the likelihood function in the usual parametric estimation. If n defines the number of subjects, $f_i(t_i)$ is the density function of the failure time T_i , $S_i(t_i) = P(T_i > t)$ is the survival function, t_i is the minimum of the exact failure time T_i , C_i is the censoring time of the i th individual and $\delta_i = I(T_i \leq C_i)$ an indicator variable which represents the failure status, then the maximum likelihood function contains two parts as below:

$$L(\beta) = \prod_{i=1}^n [f_i(t_i)]^{\delta_i} \times [S_i(t_i)]^{(1-\delta_i)} \quad (3.14)$$

If the event occurred for the individual then, $\delta_i = 1$, and the second term will be 1 and we only have $f_i(t_i)$ for it. On the other hand, if the event does not happen then we only

care about the second part which will give the probability that an individual survives over time. Also, we know that

$$h_i(t_i) = \frac{f_i(t_i)}{S_i(t_i)} \quad (3.15)$$

This is defined as hazard function at time t_i , this changes (3.13) to

$$\begin{aligned} L(\beta) &= \prod_{i=1}^n [h_i(t_i) S_i(t_i)]^{\delta_i} \times [S_i(t_i)]^{(1-\delta_i)} \\ &= \prod_{i=1}^n [h_i(t_i)]^{\delta_i} \times S_i(t_i) \end{aligned} \quad (3.16)$$

The log likelihood function, $l(\beta)$, can then be written as:

$$l(\beta) = \sum_{i=1}^n \delta_i \times \log[h_i(t_i)] + \sum_{i=1}^n \log(S_i(t_i)) \quad (3.17)$$

Based on Cox-hazard proportional model in equation (3.12), a partial likelihood can be constructed from the data as follows:

$$I(\beta) = \prod_{i:\delta_i=1} \frac{\theta_i}{\sum_{j:t_j \geq t_i} \theta_j} \quad (3.18)$$

where $\theta_i = e^{X_i \beta}$ and $(X_1, X_2, \dots, X_n)$ are the covariate vectors for the n independently selected individuals in the data set. By solving $\frac{\partial L(\beta)}{\partial \beta} = 0$, then the covariate coefficients can be obtained. To get the baseline hazard function in the full-likelihood function, β should be replaced by β' where $h_o(t_i)$ is evaluated as:

$$h_0(t_i) = \frac{1}{\sum_{j \in (t_i)} \theta_j} \quad (3.19)$$

where t_i is the time ordered for event occurrence such that $t_1 < t_2 < \dots < t_n$. The Cox proportional hazard model requires that the covariates are multiplicatively related to the hazard. In the simplest case, the exact effect of the covariates on the life-time depends on the type of $h(t)$. The Cox partial-likelihood, given in (3.19), is obtained by

using the Breslow's estimate of the baseline hazard function, substituting into the full likelihood and then observing that the outcome is a product of two factors. The first factor being the partial likelihood in which the baseline hazard cancels out. The second factor is independent of the regression coefficients and depends only on the data through the censoring pattern. The result of covariates estimated by any proportional hazards model can thus be reported using hazard ratio (Ameri, 2015).

3.3.2 Time-Dependent Cox Model (TD-Cox)

The Cox proportional hazard model makes an assumption that covariates are independent of time. In other words, when covariates are not affected over time or when data set is only collected at one time point, it is appropriate to use static variables to explain the outcome. On the other hand, there are many situations (such as the beneficiary dropouts retention problem) where the covariates change over time, and the above assumption does not hold. Thus, it is better to use time-dependent covariates which will potentially result in more precise estimates of the outcomes. Consequently, we can define time-dependent variables that can change in value over the course of the observation period. Variables such as body weight, income, marital status or number of years beneficiary spent in the programme are few examples of attributes that could vary over time. One way to look at the time varying covariates is to hold the values of such variables fixed at a certain point in time, say baseline, but to have an accurate analysis, the best way is to change variables over time. To extend the hazard function to include variables that change over time, for each time varying covariate in the model, we can represent it as a function of t . Thus, Cox proportional hazard model can be written as;

$$h(t|Z(t)) = h_0(t)e^{\beta_1 Z_1(t) + \dots + \beta_q Z_q(t)} = h_0(t)e^{(Z(t)\beta')} \quad (3.20)$$

where $Z_1(t), \dots, Z_q(t)$ are the covariates, β_i are the coefficients, h_0 is the baseline hazard function and t the time. So, this function means that the hazard at time t is dependent on the value of Z at time t . Extensions to time varying attributes can be incorporated using

the counting process formulation. Essentially, in the counting process, data are expanded from one record-per-subject to one record-per interval between each event time for each subject. The covariates information need. to be updated and available at these times, but not in between (Tarone and Ware, 1977).

KNUST



Chapter 4

Data Analysis and Results

4.1 Introduction

This chapter presents the data analysis and discussion of results. The results are presented under the following sub-headings: retention (survival) of beneficiaries, Cox Hazard proportional (regression) model and discussion. The study used secondary data from Compassion International database called 'Compassion Connect'. Data collected include the age, gender, caregiver educational background, location of beneficiaries and religion. They were categorised into various levels as indicated in the Table 4.1.

Table 4.1: Categorisation of data collected

Variables	Definition	Measurement
Dependent		
Indicator	Status of beneficiaries	0 = censored 1 = not censored
Independent		
Age	Age of beneficiary when he/she was enrolled	1 = 3-9 2 = 10-15 3 = 16-22
Edugrade	educational grade of beneficiary	1 = JHS 2 = lower primary 3 = upper primary 4 = Secondary
Religion	Religious affiliation	1 = Orthodox 2 = Pentecostal 3 = Catholic 4 = Islam
Caregiveroccu	occupation of beneficiary guardian	
Location	cluster where beneficiary lives	1 = Kasoa 2 = Breman 3 = Swedru

4.2 Retention (Survival) of Beneficiaries

4.2.1 Cluster Analysis

The survival function plots are presented in Figure 4.1 which show the survival distribution between beneficiary's clusters. It is observed that as a beneficiary number of years spent on the programme increases, the probability of surviving (censoring) decreases. Again, the figure shows that, there is an interaction between survival distributions among beneficiary clusters. However, the probability of survival is lower for Breman than other clusters (Kasoa and Swedru). This means that beneficiaries who are not close to Accra are less likely to survive on the programme. The inference may be that exposure in the hinterlands of which Breman is one is low. They have less information on benefits and opportunities for higher education and other life opportunities. Actually, they have no significant basis for comparison which can push them to stay on the programme. However, those from Swedru and Kasoa have the tendency of getting spillover effects from Accra easily since distance between these clusters and Accra is short compared to Breman.

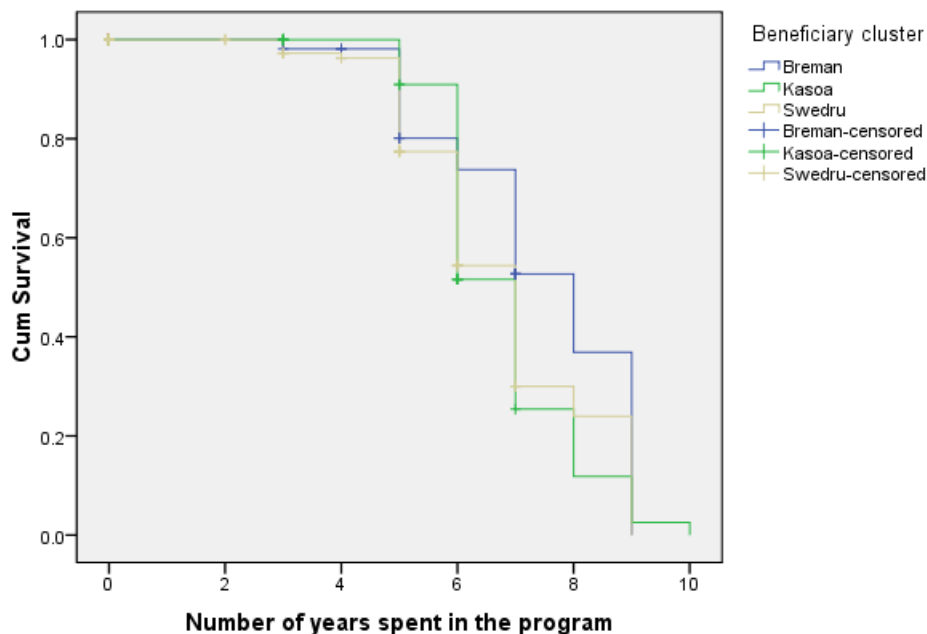


Figure 4.1: Survival function for beneficiary cluster

Table 4.2 below reports an estimated average time for a beneficiary on the (Compassion International) programme drop out in the various clusters. The table shows that average beneficiary spends about 7 years on the programme. However, cluster analysis shows some variation on the dropout time. While beneficiaries from Breman spend almost six years, six months, those from Kasoa spend seven years four months and Swedru spend six years eight months.

Table 4.2: Means for Survival Time

Beneficiary cluster	Mean			
	Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Breman	6.597	0.231	6.943	7.850
Kasoa	7.397	0.107	6.613	7.034
Swedru	6.792	0.160	6.478	7.106
Overall	6.909	0.087	6.739	7.080

4.2.2 Log Rank Test

To check whether there is difference in overall survival distribution between clusters, a log rank test was run to test the null hypothesis that there is no difference in the overall survival distributions between the three clusters (Breman, Kasoa and Swedru). Log Rank Test was performed at 5% level of significant. Therefore, if the probability value p ($p < 0.05$) is less than 0.05 implies non-acceptance of the null hypothesis. From

Table 4.3, Chi-square value of 7.580 at degree of freedom of two (2) was significant since the probability value of 0.023 was even less than the benchmark probability of 5%. Therefore, it is concluded that there is a significant evidence of a difference in survival times for the clusters.

Table 4.3: Overall Comparisons

Test	Chi-Square	df	p-value
Log Rank (Mantel-Cox)	7.580	2	0.023
Breslow (Generalized Wilcoxon)	6.123	2	0.047
Tarone-Ware	7.109	2	0.029

Table 4.4 below shows the total number of dropouts and number of censored (survived) in each cluster with percentages. The total number of beneficiaries in each cluster is shown in column 2, number of dropouts in column 3 and column 4 indicates number of survival. Breman cluster has a total number of 385 beneficiaries with censor of 88.8%, Kasoa had 1246 beneficiaries with a censor of 90% whiles Swedru had 1086 beneficiaries with 91.5%. Given survival rate, Breman cluster has a lower survival rate as indicated by the percentage in column five in Table 4.3. The percentage censoring shows that though Breman cluster had smaller number of beneficiaries, it had the highest rate of possible dropouts of 11%. Probably, inferiority complex and inappropriate level of comparison may account for this difference in survival rate on the programme.

Table 4.4: Case Processing Summary of clusters

Beneficiary cluster	Total N	N_e of Events	Censored	
			N_c	Percent
Breman	385	43	342	88.8%
Kasoa	1246	125	1121	90.0%
Swedru	1086	92	994	91.5%
Overall	2717	260	2457	90.4%

where N is the total number of beneficiaries enrolled, N_e the number who dropped out and N_c the number who did not drop out during the duration of this study

4.2.3 Gender Analysis

Gender issues have taken central stage in almost all aspects of socio-economic development in every nation. Compassion International also considers it as one of the prime objectives of its programme. Therefore, gender based analysis is considered in this study.

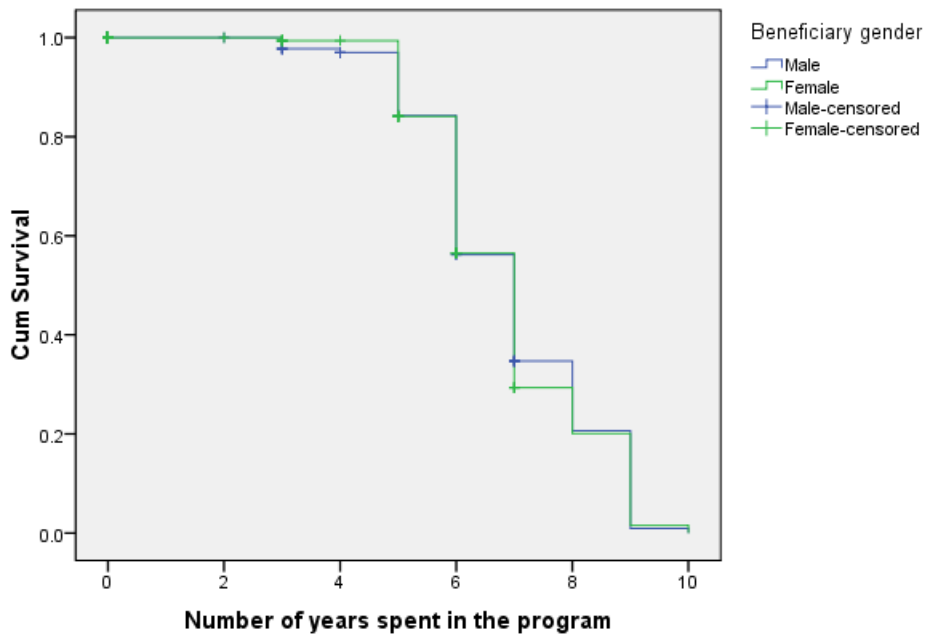


Figure 4.2: Survival functions beneficiary gender

Figure 4.2 presents survival function plot based on gender. The probabilities of surviving by males and females interact so it is difficult to tell who has lower probability to survive. However, at points where there is no interaction, males appear to have lower survival probability meaning that they are less likely to survive than females in the early periods of the programme. Females in later years may become more promiscuous which results in teenage pregnancy hence dropout from the programme.

4.2.4 Survival Rate

Table 4.2 indicates the number of males and female on the programme, where N is the total number of beneficiaries enrolled, N_e the number who dropped out and N_c the number who did not drop out during the duration of this study. The total number of observations was 2,717, made up of 1,314 males and 1,403 females. The 1,196 males survived representing 91.0% and 1,261 females survived (censored) representing 89.9%. Therefore, it can be concluded that females are slightly likely to drop out from the programme more than their male counterparts.

Table 4.5: Case Processing Summary of gender

Beneficiary Gender	Total N	N_e of Events	Censored	
			N_c	Percent
Male	1314	118	1196	91.0%
Female	1403	142	1261	89.9%
Overall	2717	260	2457	90.4%

Table 4.6 below indicates estimated average number of years a beneficiary will stay on the programme. It indicates that averagely, males are more likely to stay longer than females. Males had mean of 6.92 with females having 6.90 years. Pregnancy and marriage maybe probably explanation for this trend.

Table 4.6: Means for Survival Time

Beneficiary gender	Mean			
	Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Male	6.915	0.134	6.652	7.177
Female	6.903	0.115	6.678	7.128
Overall	6.909	0.087	6.739	7.080

The differences in gender survival rates are also tested using a chi-square test based on log Rank test. The null hypothesis being there is no difference between means at 5% level of significance is tested. The test results as shown in Table 4.7 indicate that the null hypothesis of no difference between means is failed to be rejected. All the three tests accept the null hypothesis. The decision rule is to reject null hypothesis if p-value is less than ($p < 0.05$). Since the probability values are all greater than 5% (0.05), there is enough evidence to conclude that there is no difference in the distributions between males and females dropouts. This is not surprising since the average year of dropouts were almost the same.

Table 4.7: Overall Comparisons

Test	Chi-Square	df	p-value
Log Rank (Mantel-Cox)	0.030	1	0.863
Breslow (Generalized Wilcoxon)	0.033	1	0.856
Tarone-Ware	0.043	1	0.835

4.3 Cox Hazard Proportional (Regression) Model

The Cox hazard proportional model would be fitted to the records obtained from Compassion International Database. This is to measure the risk of dropout by the individuals enrolled in the programme between the time they were registered to expected completion period.

4.3.1 Cox Hazard Proportional Model for Gender

Table 4.8, presents results of the cox regression model for gender of testing whether dropout is dependent of gender at 5% level of significance. The results that gender is not significant ($p\text{-value} = 0.896 > 0.05$), which means for the beneficiary to dropout from compassion international programme does not depend on gender. Again, the risk associated with dropout dependent on gender is 1.02 lower.

Table 4.8: Variables in the Equation for Gender

	B	SE	Wald	df	p-value	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
Gender	-0.016	0.125	0.017	1	0.896	0.984	0.771	1.256

Figure 4.3 below depicts the survival function plot of the number of years respondents stay in the programme. From the plot, it can be seen that beneficiaries are able to stay in the programme in the initial stages but dropout in the latter stages. This can also be seen in Figure 4.4 which shows the survival between male and female individuals registered in the programme. A high proportion of individuals are able to stay in the designed programme in the first two years before dropping out.

Figures 4.5 & 4.6 are plots of the hazard functions. These shows the risk of an individual leaving the programme given that he/she has stayed up to time t . It measures the risk of dropping out within a short time period. The greater the hazard between any two time periods, the greater the risk of an individual dropping out from the programme.

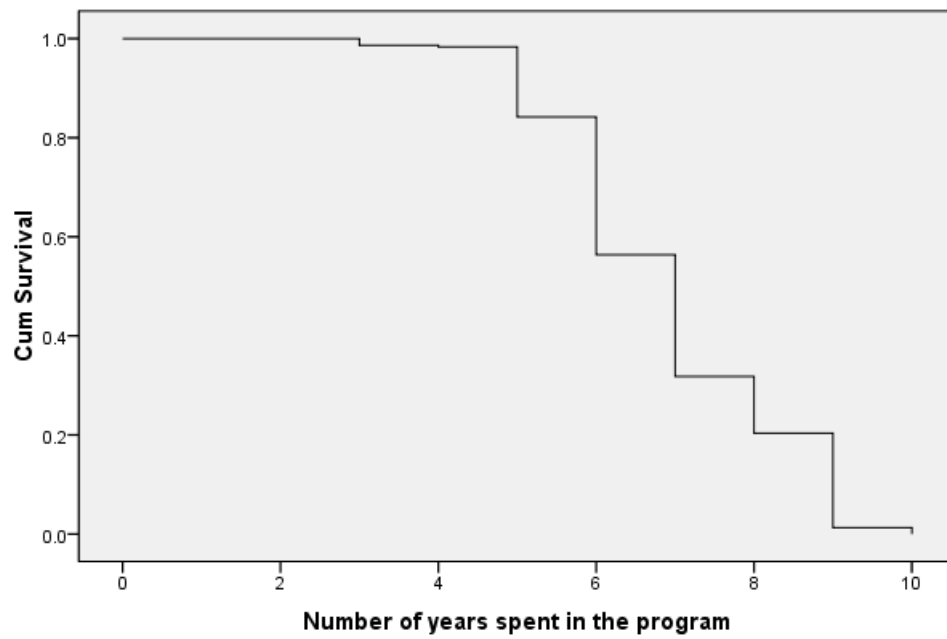


Figure 4.3: Survival functions at mean of covariates

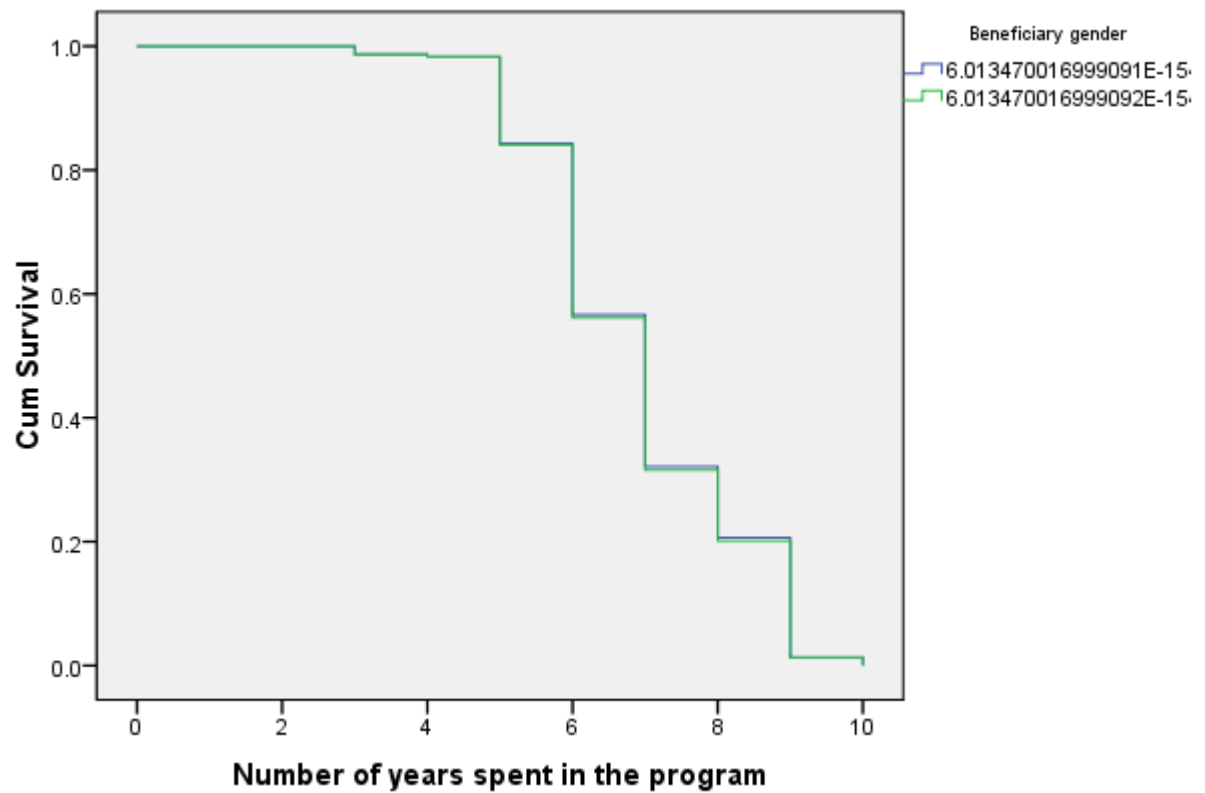


Figure 4.4: Survival function for patterns 1 - 2

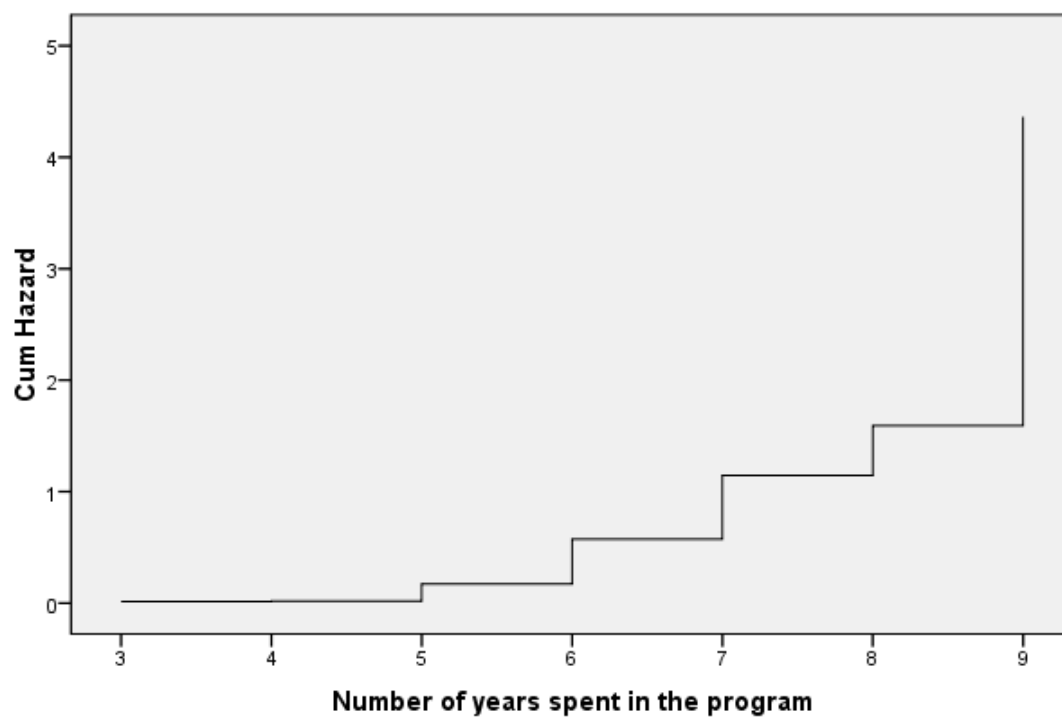


Figure 4.5: Hazard function at mean of covariates

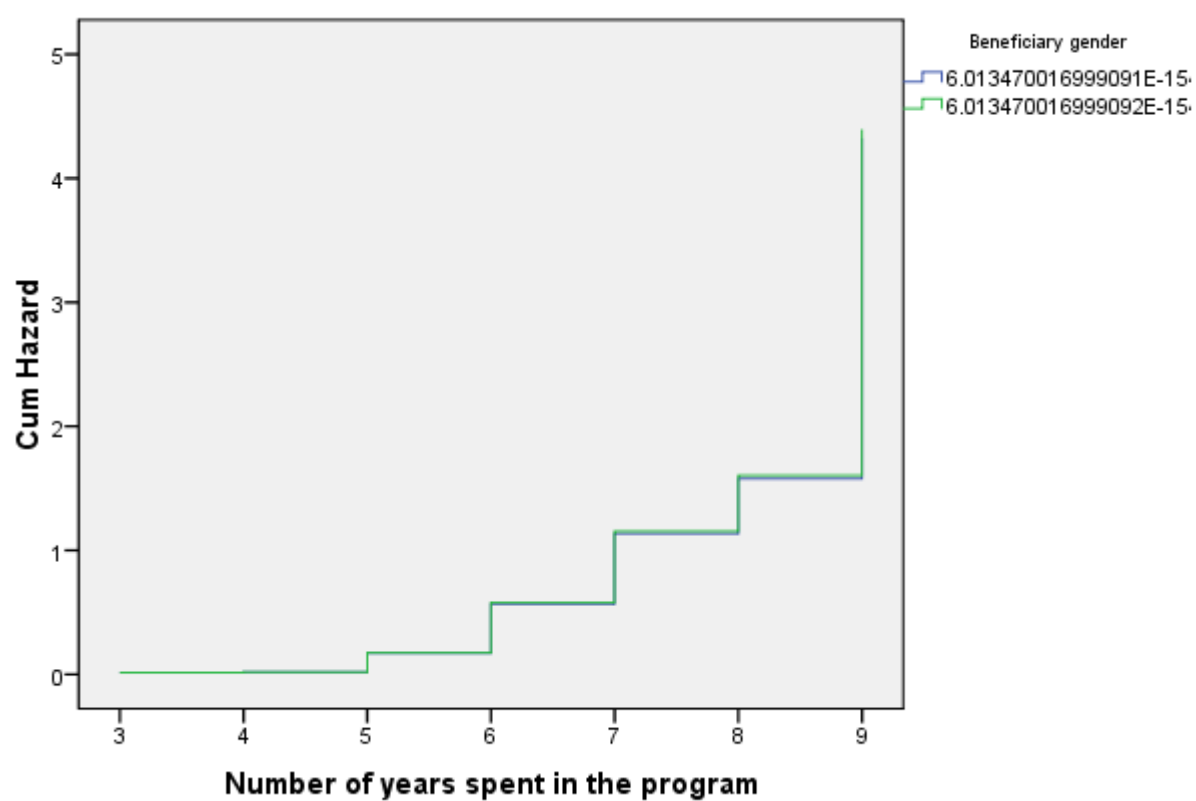


Figure 4.6: Hazard function for patterns 1 - 2

4.3.2 Combined Cox Hazard Proportional Model for Predictors.

The Cox hazard proportional model is used to express the link between the hazard and other set of variables that could explain the events (dropouts) from the programme. Results from the Cox regression model are presented in Tables 5.1, 5.2, and 5.3 in Appendix A. Table 5.1 describes the data distribution. It shows that there were 260 dropouts with 38 being rescued or reinstated. This shows that programme officers are doing well in terms of monitoring and feedback. The quick response of supervisors indicates the effectiveness of Compassion International's outreach programme. There was no missing values and negative time. Therefore, the number of cases remain the same throughout the programme since the earliest event in a stratum does not change.

Table 5.2 displays the mean value of each predictor variable in column 2 and a pattern for each level in column 3.

The Omnibus Tests shown in Table 5.4 explores whether the predicted probabilities are the same as the observed probabilities. An overall goodness of fit of the model is good since the test result is insignificant. This model produced a significant difference between the predicted and observed probabilities indicating a good model fit.

Table 4.9, presents results of the fitted Cox regression model of other predictor variables. The last column captured $\exp(b)$ which represents the hazard rate, the coefficient (b) column represents coefficients of the variables and p-value column indicates significant values of the individual variables. The Cox regression model included the following covariates: location (clusters), age, educational grade, religion and caregiver occupation. The p-value column shows that with exception of location2, age1 and edugrade1 all other variables are significant based on benchmark standard significance level of 0.05. The insignificance of location2, age1 and edugrade1 means that there is no evidence of a greater risk of beneficiaries on the Compassion international dropping out. The effect of age2 is statistically significant which means that for each additional year of beneficiaries, the hazard decreases by $(1 - 0.931) * 100 = 6.9\%$. Basing explanation on the reference category of 16-22 years, means that there is no evidence of a difference between 3-9 year

Table 4.9: Cox Hazard model of dropouts predictors

Variables	Coefficient (b)	SE	t-value	Sig	Exp(b)
Location2	0.103	0.522	0.197	0.844	1.108
Location3	0.906*	0.457	1.982	0.047	2.475
Age1	0.183	0.309	0.592	0.553	1.201
Age2	-0.071*	0.036	1.972	0.049	0.931
Edugrade1	-0.002	0.009	0.222	0.840	0.998
Edugrade2	0.539**	0.139	3.878	0.000	1.709
Orthodox Christian	0.583**	0.143	4.077	0.000	1.791
Pentecostal	1.284**	0.143	8.979	0.000	3.611
Caregiveroccu	0.151**	0.021	7.191	0.000	1.163

olds and group of 16-22 year olds. However, there is a difference between the age group of 3-9 and 10-15 years. After controlling for gender and other covariates, the hazard for 10-15-year old is 0.931 times that of above of 16-22 years.

For educational grade, the reference category is above JHS compared to lower primary (edugrade1) and upper primary (edugrade2). Educational grade one (Edugrade1) means that there is no evidence of a greater risk of beneficiaries in lower primary on the Compassion international dropping out. One unit increases in educational grade at lower primary reduces the probability of dropping out by 0.2%. For educational grade two (Educagrade2), there is evidence of difference between above JHS and upper primary. This means that one year promotion within upper primary increases the greater risk of beneficiaries dropping out from the programme by about 71%.

Religious affiliation in Ghana plays crucial role in adherent to associations. Religion was considered in the study and categorized into orthodox Christians, Pentecostals, Islam with Islam being the reference. The results from Table 5.2 shows that there is evidence of

a difference between the religions. Orthodox Christians differ from Islam and has hazard odds of 7.9 times that of Islam and Pentecostals also has 6.1 times of Islam.

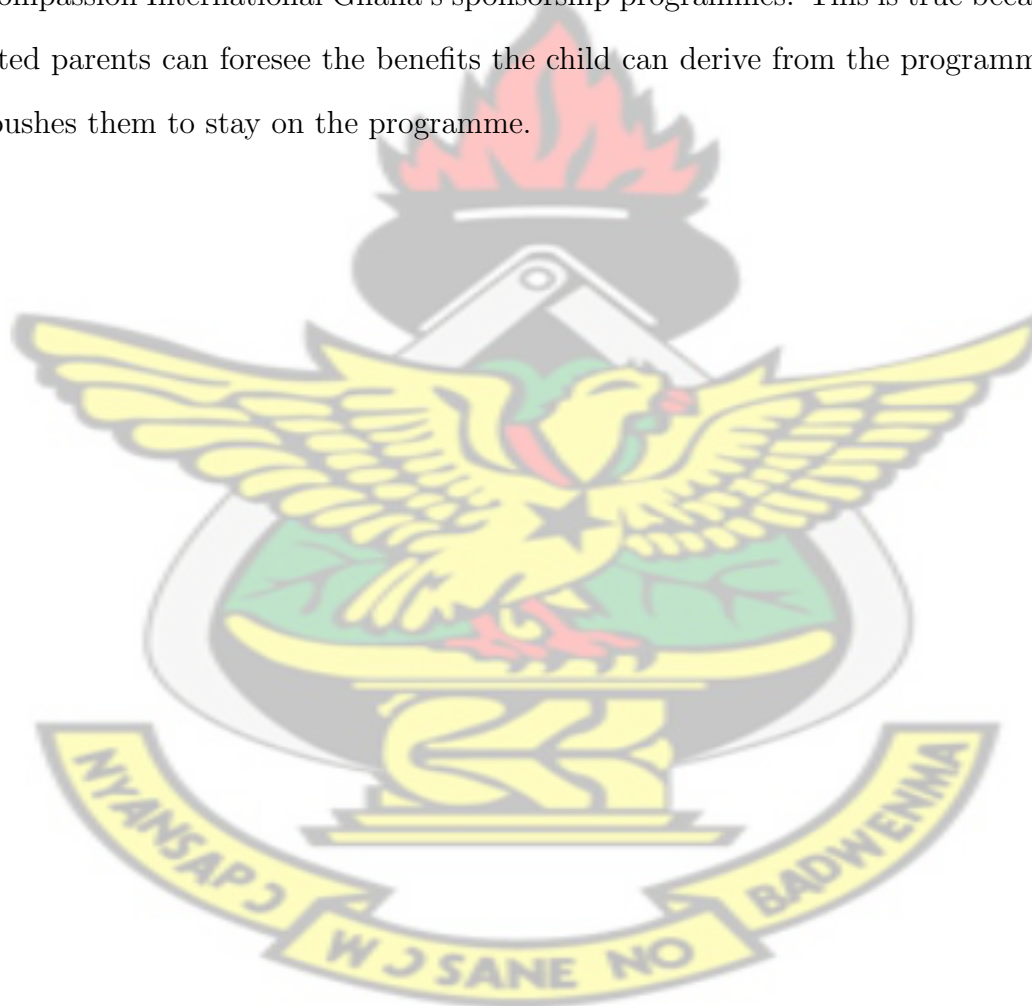
The $\text{Exp}(B)$ represents the hazard rate. The value of 2.475 for location3 means that the risk of beneficiary dropping out from the programme is 0.40 times less than those in location1. This low risk of dropping out reveals closeness of a location to the capital city and those far away from it. Location1 is Kasoa cluster. Those in location one has high exposure to modern forms of entertainment and other kinds of enjoyment. The age two has a value of 0.931 and the risk of average age beneficiary dropping out from the programme is 1.074 less than those who are older than 15 years.

4.4 Discussion of Results

From Figure 4.1, it shows that, the survival probability is lower for Breaman than other clusters (Kasoa, Swedru). This supports the study of Saeed et al (2016), whose results reveal that beneficiaries are less likely to survive in programme due to their location. Again, Okumu et al. (2008), also found out that, the survival of students diminishes as they move from urban areas to the rural areas. It implies that, there is relationship between location and survival of beneficiary. This can be possible since, beneficiaries in urban areas have access to social amenities, educated people to mentors them, and those who have benefited from the programme. All these push them to stay on the programmes as compare to their counterparts in the rural areas.

Per the gender analysis, it was concluded that, females are slightly likely to drop out from the programme more than their male counterparts. Holmes (2003) and UNICEF (1993) studies confirmed this result that female beneficiaries have less survival rate on sponsorship programmes than male counterpart, Saeed et al., (2016) also observed that, dropout rate was 3.24% for males and 7.55% for females in one of their research papers. This is true because of early marriage and teenage pregnancy. Some of females also feel shy when their woman-hood is developing because their counterparts will laugh at them. The cox regression analysis shows that, the probability of beneficiaries in lower primary dropping from programme is less as compare to beneficiaries in upper primary. This is

so because most caregivers have time for children in that grade because of their age. At that stage, caregivers know they cannot do things on their own, so they must help them out but as they are entering upper primary, the caregivers believe that they can do things on their own, so they really do not bother much to follow up on what they do. These can be linked to the educational background of the caregivers. Choudhury (2006) observed that there is a significant direct relationship between father's educational level and student's survival in programmes. The study concluded that a step above the level of a father's education leads to an increase in his ward's survival on educational programmes like Compassion International Ghana's sponsorship programmes. This is true because the educated parents can foresee the benefits the child can derive from the programmes and that pushes them to stay on the programme.



Chapter 5

Conclusion and Recommendations

5.1 Summary of Findings

The purpose of the study was to examine the determinants of beneficiary dropout (exit) from Compassion International Ghana programme and to achieve this, 2717 beneficiaries from Swedru, Breman and Kasoa clusters on the database of Compassion International for a period of ten years were taken. The following are the summary of the findings of the data analysis and observations were made:

- Location/community (rural/urban) influences dropout rates. Those from the urban communities are likely to stay on the programme more than the rural communities. This is because urban dwellers have nothing to do than going to work or school, so people are forced to get one of these urban jobs; either work or school. However, when you are young, you can only get job if and only if you have certificate, therefore children in urban communities try always to be in school in order to get certificate to work with in the future but those in the rural areas have alternatives' such as farming, hunting and playing with peers in the village. Again, for urban dwellers, because they have access to educated people, those who have gone through the programme and succeeded, and have access to social amenities, they serve as source of motivation to the beneficiaries and that compel them to stay on the programme but for rural areas their source of motivation is the community members which most of them are farmers so end up dropping out and join their parents in farming. Breman cluster is more likely to drop-out from the programme. This cluster is far away from the regional capital as well as national capital so most of the beneficiaries are limited moreover, they face low level of social competition and any benchmark for comparison. Most of them do not see the importance of the programme especially,

schooling. They have not come across with beneficiaries who have completed programme successfully to serve as motivation for them and their parents (caregivers). Parents from rural areas take along children when going to the farm because they serve as source of labour, making the beneficiaries not interested in attending programme. Rather, they prefer taking the children (beneficiaries) along when going to the farm.

- On gender basis, females are likely to dropout earlier than the male counterparts. Females' biological development is faster than their male counterpart hence they mature faster than males. Mature females attract teasing from their colleagues and as a result they dropout. Besides, irresponsible men also capitalised on the ignorance of these innocent girls and lure them with bait such as provisions, television set and other funny items like mobile phones which result in unexpected teenage pregnancy, hence dropout from the programme. Though, their male counterparts can equally impregnate them or other girls outside the programme, males' do not directly carry the burden of such unexpected pregnancy. In such situation, the girl carries the pregnancy for the specified nine months plus at least two years before the girl involved could return to school again if only she is ready and determined.
- Ages - Older beneficiaries have high probability of staying on the programme. Beneficiaries on the programme, old and mature stay on the programme till they finish because they may have encountered one or two lives difficult, so they see the programme to be the greatest opportunity to turn life round. It is their second chance in life and therefore, no stone is left unturned to amend mistakes made previously. Also these older people know what they are after so they make better use of the programme opportunity. Again, national programmes, exchange programmes, camping and so on, they might have come across with those who have gone through the programme and succeeded, so it serves as a source of inspiration to them and it force them to stay on the programme.

5.2 Conclusion

Survival of beneficiaries on the Compassion International Ghana sponsorship programme is the life blood of Compassion International Ghana. Hence the beneficiaries must be on the programme until the programme ends. Survival analysis was conducted to examine the determinants of a beneficiary dropout from Compassion International Ghana programme. It was established throughout the study that, clusters in rural communities are likely to have a high dropout rates. Also, clusters of more young beneficiaries (upper and lower primary) are likely to have high dropout rates. In addition, cluster of more females are likely to experience high dropout rates. Therefore, clusters in these categories should be given more attention so that, beneficiaries can survive until the programme ends.

5.3 Recommendations

It is recommended that the case study institution - Compassion International Ghana - should have special supervision for female beneficiaries. This is because, the findings indicate that, innocent females are being attracted by provisions, televisions, mobile phones and so on. They should be educated more about the implications of these. Moreover, the programme implementers should try to provide some basic needs like sanitary pads, under wears, and other basic needs to discourage the rural females from following irresponsible male counterparts. Again, there should be a more education on impact of teenage pregnancy on their future life. Breman cluster should be given special attention to reduce the number of drop-outs. Because they are far from capitals or cities which serve as benchmark for comparison, more forum about the importance of the programme should be advocated. Those who have gone through the programme and succeeded can be invited to talk to them about the programme and that will motivate them to stay on the programme. There can be a career day celebration where they will have opportunity to meet people with different careers, to serve as their mentors. Moreover, there can be exchange programmes where beneficiaries from Bremen can visit different clusters to learn from them. All these will inform their decision-making in the future. Finally, in terms of

distribution of educational materials and items, they should be considered first to serve as source of motivation for them to stay on the programme. Forum and workshops should be periodically organized in the rural communities for parents and beneficiaries to know the future prospects of the programme. This is very important because when the beneficiaries and caregiver are aware of the future prospects of the programme, it encourages them to stay on the programme. The forum can include those who have benefited from the Compassion International programme before. Currently, some of the beneficiaries have finished tertiary education and other are working, they can bring them on board to share their experiences with them.



REFERENCES

- Adamson, P. (1996). *The Progress of Nations, 1996: The Nations of the World Ranked According to Their Achievements in Child Health, Nutrition, Education, Family Planning, and Progress for Women*. ERIC.
- Ameri, S. (2015). Survival analysis approach for early prediction of student dropout.
- Bland, J. M. and Altman, D. G. (2004). The logrank test. *British Medical Journal*, 328(7447):1073.
- Choudhury, A. (2006). Revisiting dropouts: Old issues, fresh perspectives. *Economic and Political Weekly*, pages 5257–5263.
- Chugh, S. (2004). *Why children drop out?: Case study of a metropolitan slum*. Bookwell.
- Compassion, International, I. (2017). Program field manual. *Compassion International Program Field Manual*, 2.5:29–30.
- Cornell, U. (2010). Statnews. *Cornell university Statistical Unit*, 78.
- Finn, J. D., Stott, M. W., and Zarichny, K. T. (1988). School performance of adolescents in juvenile court. *Urban Education*, 23(2):150–161.
- Government (2011). Ghana school feeding programme annual operating plan government of the netherlands (2011) republic of ghana.
- Gupta, S. and Singh, S. K. (2016). Economic security among rural women through self help groups: An analytical study. *Indian Research Journal of Extension Education*, 12(2):117–122.
- Hanushek, E. A. and Wößmann, L. (2007). *The role of education quality for economic growth*. The World Bank.
- Hartnagel, T. F. and Krahn, H. (1989). High school dropouts, labor market success, and criminal behavior. *Youth & Society*, 20(4):416–444.

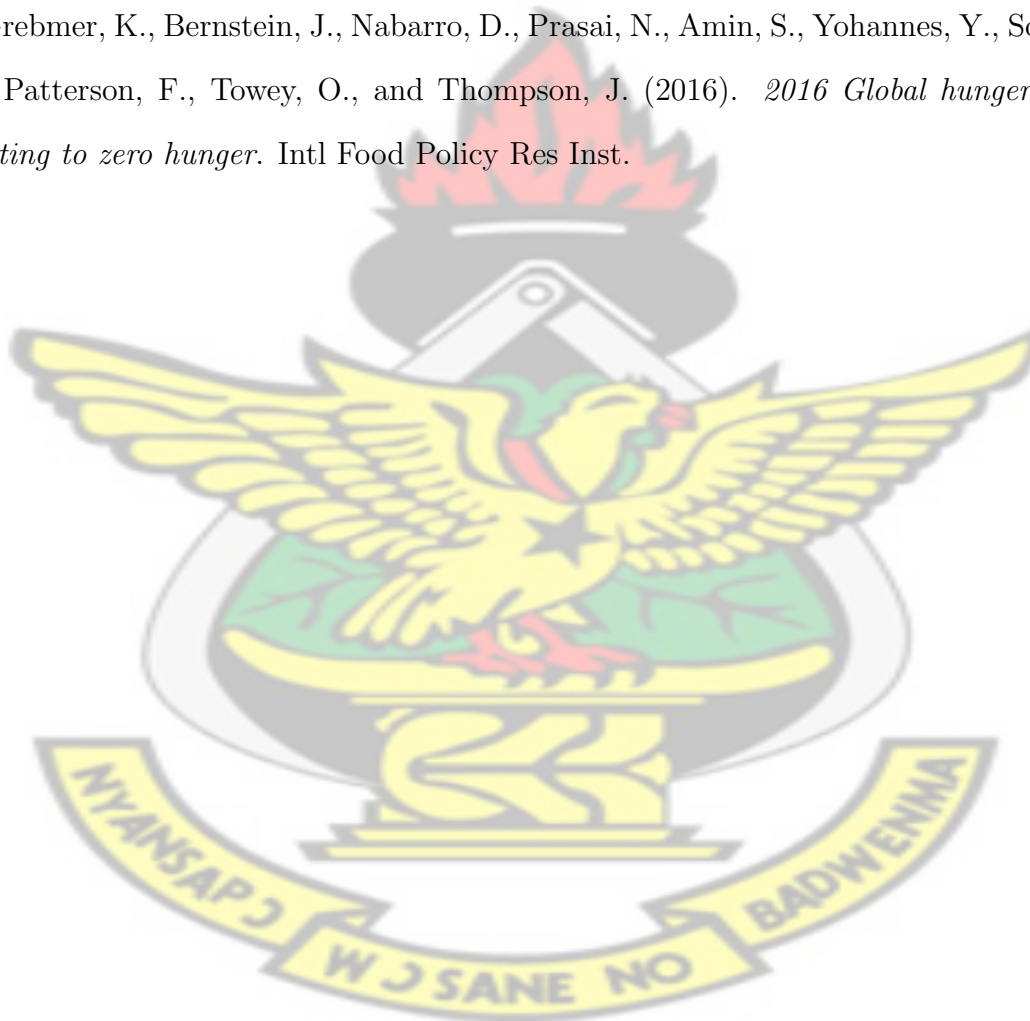
- Holmes, J. (2003). Measuring the determinants of school completion in pakistan: analysis of censoring and selection bias. *Economics of Education Review*, 22(3):249–264.
- Hosmer Jr, D. W. and Lemeshow, S. (1999). Applied survival analysis: regression modelling of time to event data (1999). *European Orthodontic Society*, pages 561–562.
- Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, 53(282):457–481.
- Khan, M. H. (2001). *Rural poverty in developing countries: implications for public policy*, volume 26. International Monetary Fund.
- Lawry, L. (2009). Guide to nongovernmental organizations for the military. *Center for Disaster and Humanitarian Assistance Medicine (CDAHMA), Summer*.
- Lazar, I., Darlington, R., Murray, H., Royce, J., Snipper, A., and Ramey, C. T. (1982). Lasting effects of early education: A report from the consortium for longitudinal studies. *Monographs of the Society for Research in child Development*, pages 1–151.
- Lewis, D. and Kanji, N. (2009). *Non-governmental organizations and development*. Routledge.
- Miles, S. (2009). Engaging with teachers’ knowledge: promoting inclusion in zambian schools. *Disability & Society*, 24(5):611–624.
- Milligan, K., Moretti, E., and Oreopoulos, P. (2004). Does education improve citizenship? evidence from the united states and the united kingdom. *Journal of public Economics*, 88(9-10):1667–1695.
- Odia, L. and Omofonmwan, S. (2007). Educational system in nigeria problems and prospects. *Journal of Social Sciences*, 14(1):86–85.
- Okumu, I. M., Nakajjo, A., and Isoke, D. (2008). Socioeconomic determinants of primary school dropout: the logistic model analysis.

Tarone, R. E. and Ware, J. (1977). On distribution-free tests for equality of survival distributions. *Biometrika*, 64(1):156–160.

Tetteh, A. (2016). *THE ROLE OF NGOs IN ENHANCING GIRL CHILD EDUCATION: THE OF WORLD VISION INTERNATIONAL IN THE NADOWLI DISTRICT OF THE UPPER WEST REGION OF GHANA*. PhD thesis.

Upchurch, D. M. and McCarthy, J. (1990). The timing of a first birth and high school completion. *American Sociological Review*, pages 224–234.

Von Grebmer, K., Bernstein, J., Nabarro, D., Prasai, N., Amin, S., Yohannes, Y., Sonntag, A., Patterson, F., Towey, O., and Thompson, J. (2016). *2016 Global hunger index: Getting to zero hunger*. Intl Food Policy Res Inst.



Appendix A

Table 5.1: Case Processing Summary

		N	Percent
Cases available in analysis	<i>Event^a</i>	260	9.6%
	Censored	38	1.4%
	Total	298	11.0%
Cases dropped	Cases with missing values	0	.0%
	Cases with negative time	0	.0%
	Censored cases before the earliest event in a stratum	2419	89.0%
	Total	2419	89.0%
Total		2717	100.0%

a. Dependent Variable: Number of years spent in the programme

Table 5.2: Covariate Means and Pattern Values

Age	17.304	17.304	17.304
Educational grade	1.000	1.000	1.000
Religion	.121	.121	.121
Caregiver Occupation	.963	.963	.963

Covariate Means and Pattern Values are reported in table 4.9 above

Table 5.3: Covariate means for Location and Gender

	Mean	Pattern	
		1	2
Location(1)	0.178	0.178	0.178
Location(2)	0.460	0.460	0.460
Gender	0.450	1.000	0 .000

Table 5.4: Omnibus Tests of Model Coefficients

-2 Log Likelihood	Overall (score)			Change From Previous Step			Change From Previous Block		
	Chi-square	df	Sig.	Chi-squared	df	Sig.	Chi-square	df	Sig.
2555.672	4.686	6	0.585	4.920	6	0.554	4.920	6	0.554

- a. Beginning Block Number 0, initial Log Likelihood function : -2 Log Likelihood : 2560.592
- b. Beginning Block Number 1. Method = enter