Effect of family structures on undergraduate degree attainment in Kenya: The 2009 Housing and Population Census.

Submitted by;

CHARLES MBUGUA MBURU

I56/64856/2010

A research project proposal submitted in partial fulfillment of the requirement of Master of Science in Social Statistics degree, College of Biological and Physical Science, School of Mathematics, Chiromo Campus, University of Nairobi.

June 2012

Declaration

This project report is my original work and has not been presented for a degree in any other university.

The paper is submitted for examination by:

Charles Mbugua Mburu

Reg No: I56/64856/2010

Signature____________________

Date 25/07/2012

Submitted to: **Mrs Anne Wang'ombe**

Submitted for examination with my approval as university supervisor:

Signature ____________________

Date 26-07-2012

Dedication

To all my lecturers, loving family, friends, colleagues at college and at work.

You all have been so instrumental thus far.

Acknowledgement

I am deeply indebted to many people in writing this work. Above all I sincerely thank the Almighty God for the gift of good health throughout the journey and His divine guidance and strengthening every step of the way. More-so, I am deeply humbled by valuable guidance from my supervisor. Not also forgetting the support I received from senior data processing manager; Kenya National Bureau of Statistics, Mr. Mutua Kakinyi. I will forever be grateful to the generous support from my brother throughout my studies.

Abstract

Numerous studies have been conducted in the area of social science trying to establish what influences education attainment for children. These studies have sort to address areas ranging from the mode of delivery during teaching to social background factors that determine education attainment of a child. In areas touching on social background factors that influence education attainment for children, Family Structure has featured prominently in many studies across different countries.

Studies on the effect of Family Structures on education attainment have however been focusing on intact versus non-intact Family Structures. This particular study strives to go further than just the two broad categories and explore the effect of the various Family Structures existing in Kenya on numbers attaining undergraduate studies.

Making use of the 2009 Kenya National Housing and Population Census, the study goes deeper to evaluate the effect of six existing Family Structures in Kenya on education attainment for usual members of a household while also considering the family size, social economic status of the household and educational level of household head.

To establish effects of Family Structures on education attainment, counts of those usual members of a household who had completed undergraduate studies from the various family structures was obtained and regressed using Poisson regression and zero-inflated Poisson models while considering and not considering family size, Social economic status of a household and highest educational level of household head.

The study finds that significant differences exists in number of those who have attained undergraduate studies from the different Family Structures in Kenya. Polygamous, Widowed and

Divorced Family Structures affected education attainment on equal measure; that is there is no significance difference in the effect of these Family Structures on education attainment. Children from Separated family structures are the most affected.

A deviation from most of previous findings is witnessed in this study because unlike most of the other studies that found children from families with both parents excelling better in education, it is not the case for this study. Effect of Never Married Family Structure on numbers attaining undergraduate studies is found to be more favorable than effect of Married monogamous Family structure (intact family). With both parents from a married monogamous struggling to coup with current harsh economic times whereby they are forced to leave their houses early and return late, children from these households seem not to be accruing the benefits of having both parents in their lives.

This study has recommended designing of special programmes by ministry of education which should include counseling units within schools aimed at mitigating effects of Family Structures on children's education attainment and particularly children from Separated.

Table of Contents

1 Chapter 1: INTRODUCTION

1.1 Background

Evolving around "social structure and personality" or "social structure and psychological well-being", various studies have been undertaken aimed at addressing inequality that has existed from generation to generation for children from different family backgrounds. Wendy Y. Carter (1999) in the study on the effects of changing family structures on higher education for black and white cohorts concluded that growing up in a non-intact family clearly has a negative effect on adult educational attainment. Uwaifo (2008) in the study; The Effects of Family Structure and Parenthood on the Academic Performance of Nigerian University Students concluded that significant differences existed between the academic performance of students from single-parent family and those from two-parent family structures.

The "pathology of matriarchy" hypothesis that came out of the Moynihan Report (1965) concluded that the absence of a father is destructive to children, particularly boys, because it means that children will lack the economic resources, role model, discipline, structure, and guidance that a father provides. Social science research has produced evidence both for and against the "pathology of matriarchy" view. Some studies using national samples show that children from single-mother families have lower attainments than children from two-biological-parent homes; Duncan and Duncan (1969), McLanahan and Sandefur (1994), while other studies, also using national samples, show that once other factors are taken into account, children from single-mother families do approximately as well as children from two-biological-parent families; Biblarz, Raftery, and Bucur (1997), McLanahan (1985). Some studies show that alternative family types—single-mother, single-father, and stepfamilies, for example—have similar, negative consequences for children(Dawson(1991)), while other studies show that children from some kinds of nontraditional families have higher attainments, on average, than children from other kinds; Amato and Keith (1991b).

Alive to the fact that the diversity in these findings is attributable to the choice of variables during a study, this particular study seeks to evaluate the effect of family structures on education attainment in the Kenyan context on the basis of 2009 Housing and Population Census. Among the studies which have attempted to link family back ground and education in Kenya is one by Wambugu (2002) which looked at the education of workers and their family background concluding that having well-educated parents is associated with great attainment in education and earnings of workers. Another study in Kenya by Claudia Buchmann (2000) looking at family background and children enrollment in schools concluded that parents' expectation of future financial help and perceptions of labor-market discrimination against women are significant determinants of enrollment. No study has so far endeavored to evaluate within the Kenyan context the effect of family structures on the numbers attaining higher education.

Considering the various variables captured during the census, the first section of this study uses Principle Component Analysis to construct an asset-based Social Economic Status (SES) index for the various households. The second section applies Poisson Regression Model and Zero-inflated Poisson Regression Model to model the SES index obtained, family size, highest educational level of household head and the number of those who have completed undergraduate studies from a particular household. These are the best model to use since the study is dealing with count data.

The main purpose for this study is to answer these questions; - (1) Are there significant differences in the number of those attaining undergraduate studies from the different family structures in Kenya in the absence and presence of other background household factors crucial for education attainment? (2) If significant differences exist between the various family structures, are children from male-headed and female-headed households affected differently. The census data offers variables that will help answer all these questions. The only limiting factor is that it does not offer variables that can aid in evaluating how the effect of the family structures has evolved with time. Religion of household head which is also an important factor that influences education attainment of children could also not be availed due to sensitivity concerns surrounding issue of religion.

This study thus strives to offer some insightful information on how advantaged or disadvantaged vast majority of Kenyan children from various family structures have been in as far as higher education attainment is concerned.

1.2 Problem Statement

The Government of Kenya is committed to the provision of equal access to quality and relevant education and training opportunities to ALL Kenyans. Towards this goal, the government has ratified and domesticated various global policy frameworks on education. The government signed Article 26 of the Universal Declaration of Human Rights (1948), consequently recognizing and committing to the right of every child to access education. The Article recognizes the intrinsic human value of education, underpinned by strong moral and legal foundations. Other international policy frameworks ratified and signed by the government include, (but are not limited to) the 1989 United Nations Convention on the Rights of the child (CRC), the 1990 African charter on the Rights and Welfare of the child, Salamanca Statement (1994), the Millennium Development Goals (MDGs) and *most importantly the Framework for Action on Special Needs Education (1999).* This paved way for looking into programmes of how children who deviate from the average and who cannot profit substantially from standard programmes without additional help can be kept abreast with the rest on matters education.

Over the years, focus on special needs has however been mainly be on disability without much attention been dedicated to the inability by most children to profit substantially from present standard learning programmes due to the effects of the different family structures they come from. Needless to say that most children are adversely psychologically affected by the kind of family structures they come, it has been a gross oversight on the side of the government to have over the years adopted a standard approach of teaching for all children who have no disability without due consideration of the kind of family structures they come from.

It is common knowledge to anyone that family background plays a crucial role in molding and determining success in all spheres of an individual's life and particularly education attainment. Thinking of an ideal family background, what immediately hits one's mind is the traditional biblical family setup where both the mother and father are there for the children. Due to prevailing economic and social factors recent trend has seen most families drifting away from this ideal family setup with most children being brought up under varying family structures. McLanahan and Sandefur (1994) point to three factors that have increased the prevalence of single mother families over the past three decades. The first two factors deal with the growing economic independence of women. First, they suggest that women's economic independence allows women to become more selective in the choices that they make with regard to marriage and divorce. Women who have their own source of income can leave a bad marriage or decide not to get married if they become pregnant. This is an illustration of one factor among many that has lead to many children in Kenya being brought up in families where both parents are not there. This state of affairs has definitely imparted negatively on the number attaining higher education because of the absence of an essential social fabric.

Much as factors like availability of resources, facilities and study materials may be pivotal in education attainment, parental support for the children plays a key role and in its absence the effect of these other factors towards education attainment may not be as much. Psychological well being and right attitude towards education by children is all nurtured by support from both parents. That is to say that effect of family structure in the context of the kind of support a child may be getting from the family overrides every other factor that could be affecting children's education attainment. According to Professor Charles Desforges with Alberto Abouchaar, in their article "The Impact of Parental Involvement, Parental Support and Family Education on Pupil Achievements and Adjustment"; parental involvement in the form of 'at-home good parenting' has a significant positive effect on children's achievement and adjustment, even after all other factors shaping attainment have been taken out of the equation. In the primary age range the impact caused by different levels of parental involvement is much bigger than differences associated with variations in the quality of schools. The scale of the impact is evident across all social classes and all ethnic groups.

There is therefore need to evaluate the effect of family structures on education attainment above all other factors. Our current education system treats all children alike regardless of the kind of family structure they come from. By so doing all children are not on level playing ground for excelling in education circles. This study strives to evaluate if the numbers attaining undergraduate studies in Kenya from the various family structures is significantly different and if so, then it should be a matter of priority for the ministry of education to consider designing a special programme alongside the mainstream curriculum, that strives to create some kind of harmony for the various children coming from the different family structures. Otherwise in the absence of any action, majority of children who in essence just need some timely counseling to unleash their full potentials will continue wallowing in isolation with our country continuing to lose otherwise would-be great personalities.

2 Chapter 2: LITERATURE REVIEW

2.1.1 Poisson Regression on education

Previous research has applied Poisson regression on matters education. Aldieri and Vinci (2009) in their study; "Number of Children and Education in Italy" used Poisson Regression Model to the number of children ever born, which is a count data and education of parents. This was a study on a sample of 1,033 families from 1997- 2005 Longitudinal Investigation on Italian Families (ILFI) dataset. Similar study; "Women's Educational Attainment and Intergenerational Patterns of Fertility Behavior in Kenya" by Rasugu (2003) used Logistic Model because of the nature of the dependent variables where two binary responses one on preference for more children and the other preference for contraceptive method were modeled against educational level of women respondents. This was a study where a sample of married and not pregnant women was taken from the Kenya Demographic Health Survey (KDHS) data making up a sample size of 4,324 women.

When it comes to count data, some authors have described Poisson Regression as the benchmark model for count data (Cameron and Trevedi 1998; Allison 1998a, 1998b; Long 1997). Poisson regression is also increasingly being used to estimate multiplicative models for other non-negative data, Manning and Mullahy (2001) and Santos Silva and Tenreyro (2006).

This study unlike other studies which have used relatively small samples, is making use of the large cross-sectional data from the 2009 Housing and Population Census, and looking at the count data of undergraduates from the various family structures. Mostly the count data concentrates on a few small discrete values, say 0, 1 and 2; skewed to the left; and intrinsically heteroskedastic with variance increasing with the mean. In fact, virtually all the data is restricted to single digits, and the mean number of events is quite low.

These features motivate the application of special methods and models for count regression. There are two ways to proceed. The first approach is a fully parametric one that completely specifies the distribution of the data, fully respecting the restriction of responses to nonnegative integer values. The second approach is a mean-variance approach, which specifies the conditional mean to be nonnegative, and specifies the conditional variance to be a function of the conditional mean.

The second approach has been adopted and applied Poisson Regression which is not without its own limitations. In practice, the conditional variance of the data is often larger or smaller than the conditional variance implied by the Poisson model; phenomena known as over-dispersion or under-dispersion, respectively. Cases of over-dispersion are most common and could be due to variability of the incidence rates that is not fully accounted for by the included covariates in the model. Prevalence of zero counts of undergraduates from households has prompted this study to also employ zero-inflated Poisson regression model and compare results from these two models.

2.1.2 Zero-inflated Poisson Regression on education data

Use of Zero-inflated Poisson Regression has previously been seen in education circles when Shin (2011), in the study titled; - "Mixed-effects and mixed-distribution models for count data with applications to educational research data" applied Zero-inflated Regression to model the outcome of reading ability of kindergarten children aged between 5 and 7. Data was collected from 461 students and the dependent variable was the count of letters read with correct pronunciation in sixty seconds time period. The data showed excessive zero scores warranting the use of ZIP Model.

For this study, excessive zeros are prevalent due to the fact that some households had no individuals who were of age to undertake undergraduate studies and in other instances those households who had individuals of age to undertake undergraduate studies had none who had actually attained undergraduate studies.

2.1.3 Theoretical Framework

When it comes to educational attainment, economic theories focus on social and economic factors in the home and in the proximate environment. Gary Becker's (1993) household production theory in addition to the human capital theory directly links household resources and investment to the educational attainment of children. This study looks at household productivity in the context of educational attainment of children based on their parental and family socio-economic factors. SES index of a household and the size of usual members of a household are thus some of the background variables which, this study has incorporated.

The household production theory is an outgrowth of two theories, the human capital theory and the theory of allocation of time. Although these two theories view education as an investment rather than consumption, the household theory takes on a narrower viewpoint on investments dealing solely with the household. Household economics considers the family as not only a consuming unit but also as a producing unit. This theory states that a combination of time and resource inputs produce different types of commodities (Becker, 1993). In order to produce what Becker calls "quality children," parents must spend time at home and devote real resources to foster an environment that promotes and provides formal education (1993). Since families differ, time and money spent on investments will vary, as will attitudes that may be conducive to children's ability and willingness to learn. Educational level and Religion of parents will definitely determine the attitude as well as quality of time and resources parents input for their children and it is against this backdrop that my study has also included educational level of household head as background variables.

2.2 Main Objective

Gain some insightful analysis of any significant discrepancies that may be there between the numbers of undergraduates coming from the various family structures in Kenya, sufficient enough to warrant coordinated interventions by the ministry of education.

The Specific Objectives

Establish if indeed significant differences exists between the numbers of children attaining higher education from the different family structures in Kenya.

Establish if children from Female- headed and Male- headed household are affected differently or would it otherwise call for different approaches for each of them when it comes to efforts of mitigating the effect of family structure.

3 Chapter 3: METHODOLOGY

3.1 Computing Social Economic Status (SES) index using Principle Component Analysis (PCA)

PCA is applied to construct a SES index for each conventional household from Peri-urban area based on the 2009 National Housing and Population Census data. The approach adopted is one by Filmer and Pritchet (1998) that employs use of factor analysis to define wealth indices for households. Using factor analysis, durable and non durable assets are considered alongside selected household characteristics that are closely related to economic well being of a household with a view of observing a few hidden "background' variables or aspects of economic well being of a household that is common to the observed variables and not directly visible. The aim is to reduce the complexity of the observed correlated household variables to a few of the hidden uncorrelated "background" variables also known as factors or components when the method of principle component is used.

Among the observed variables of a household considered are livestock ownership, various assets in the house, materials of construction of the house, household amenities like waste materials disposal methods and main sources of water, main economic activity of the household head and the highest level of education completed by the household head.

The first step is to re-code selected categorical variables of household characteristics and household head attributes into binary variables of Yes=1 and No=0 to avoid the categories being converted into a quantitative scale that is not helpful. This will distinguish between the presence or absence of an asset, characteristic or an attribute. Next is a descriptive analysis of all these binary variables giving their means, standard deviation and frequency. This helps shed light on which variables to combine or eliminate based on their frequencies.

PCA is a multivariate statistical technique which reduces the initial set of say n correlated household variables into smaller number of 'dimensions' of say m uncorrelated indices, where each indice is a linear weighted combination of the initial household variables. These household variables are combined such that the maximum contribution to a given aspect of SES variance between households is extracted from the variables. This variance is removed and then a second linear combination which again explains another maximum proportion of remaining variance of SES aspects of a households, and so on. This is the principle axis method and results in orthogonal (uncorrelated) factors as follows;-

$$\begin{array}{ccccccccc} PC_1 & = & a_{11}X_1 + a_{12}X_2 + & a_{13}X_3 + a_{13}X_4 & + & \cdots & a_{1n}X_n \\ \vdots & \vdots & \vdots \quad \vdots & \vdots & \vdots & \vdots & \vdots \\ PC_m & = & a_{m1}X_1 + a_{m2}X_2 + & a_{m3}X_3 + \quad a_{m4}X_4 + & & \cdots & a_{mn}X_n \end{array}$$

Where a_{mn} represent the weight for the m^{th} principle component and the n^{th} household variable.

The system of equations is expressed as;-

PC = Ax Where PC = $(PC_1 \ldots\ldots\ldots\ldots, PC_m)$ are m Principle Components.

A=Matrix of coefficients of the assigned weights and $X = X_1 \ldots\ldots\ldots, X_n$ are selected household characteristics under consideration.

The original data is not standardized and therefore the weights are the eigenvectors of the correlation matrix of weights. The variance (λ) for each principal component is given by the Eigen value of the corresponding eigenvector. The components are ordered so that the first component (PC1) explains the largest possible amount of variations in the original data, subject to the constraint that the sum of the squared weights is equal to one i.e. $a_{11}^2 + a_{12}^2 + \cdots + a_{1n}^2 = 1$

The second component (PC_2) is completely uncorrelated with the first component, subject to the same constraint. Subsequent components are uncorrelated with previous components. Each component captures an additional dimension in the data, while explaining smaller and smaller proportions of the variation of the original variables. The higher the degree of correlation among the original variables in the data, the fewer components required to capture common information.

SES index of a particular household is derived from the PCA output Analysis by taking the first principle component factor scores and constructing a dependent variable for each household(E_i) which has a mean equal to zero and a standard deviation equal to one as follows;-

$$E_i = \sum_{i}^{n} f_i \frac{(a_{ij} - a_i)}{s_i}$$

Where

E_i Is the social economic index for a household (j= 1, 2 ….n).

f_i Is the scoring factor for each observed household variable (j= 1, 2 ….n).

a_{ij} Is the i^{th} variable for the j^{th} household (i, j= 1, 2 ....n).

a_i Is the mean of the i^{th} household variable (i= 1, 2 ....n).

s_i Is the standard deviation of the household variable (i= 1, 2 ....n).

Positive factor score is associated with higher SES and negative factor score is associated with lower SES.

3.2 Poisson Regression Modeling

When it comes to count data, we have two categories of counts; Counts in space and count over a specified interval of time. For this study, counts are in space during the census exercise for those indicated to have completed undergraduate level of education and were usual members of a household.

A categorical variable is created from the constructed household SES such that I have Index

$E_i = \begin{cases} 1 = Poor \\ 2 = Middle\ class \\ 3 = Rich \end{cases}$ to be modeled amidst other socio-economic factors of a household in estimating the effect of family structures on numbers attaining undergraduate studies.

This study is concerned with evaluating if there is existing differences in the number of those who have completed undergraduate studies from different households taking into account compositional changes in family structures and other social background parameters that are deemed to influence education attainment.

Letting Y_i= count of Undergraduates from a household with a set of characteristic,X_i, be independent Poisson variable with mean = λ_i Probability of then observing an undergraduate; Y_i from a household is given by:

$$P[Y_i = y] = e^{-\lambda}\frac{\lambda^y}{y!}, \qquad y = 0,1,2, ...$$

Where E[y]; the expected value of Y_i = V(y); variance of Y_i = λ. This is the equi-dispersion property assumed of a Poisson distribution.

In fitting my Poisson Regression model, data is used as $Y = \sum Y_i$; aggregate of undergraduate counts per n_i households with characteristics X_i. The idea is to model the mean rate of undergraduate counts in a group of households as a function of household characteristics.

To do this, an additional important property of the Poisson distribution which stipulates that the sum of independent Poisson random variables is also Poisson is used. Specially, if Y_1 and Y_2 are independent with $Y_i \sim P(\lambda_i)$ for i = 1; 2 then

$Y_1 + Y_2 \sim P(\lambda_1 + \lambda_2)$:

Then:

$Y \sim P(\lambda)$

Where, λ is the expected mean rate of undergraduate counts in a group of households with characteristics of study.

Since Y depends on a set of household characteristics; X_i some observed and some unobserved, then a simple linear model of the form:

$\lambda_i = X'_i\beta$ can be expressed.

Thinking of this as a generalized linear model, the Poisson distribution of the expected mean of undergraduates counts which is the stochastic part, is related to the deterministic part of linear predictors for household characteristics through a link function that can be defined as:

$\eta = \sum X'_i\beta$

Writing Poisson distribution as an exponential family:

$$f(y, \lambda) = e^{-\lambda}\frac{\lambda^y}{y!}$$

$$= \exp(-\lambda)\frac{1}{y!}\exp(y log\lambda)$$

The natural parameter is $log\lambda$, so the canonical link function is the log link, $\eta = log\lambda$.

This is a monotonic differentiable function that ensures estimates of $\lambda \in [0, \infty)$.

Thus, the generalized linear model can be considered as an additive log-linear model:

$log\lambda_i = \sum X'_i\beta$ (1)

The expected mean of undergraduate counts is linked to household characteristics through the log-link function which is the natural logarithm. This is contrary to what happens with normal linear models where it is the mean itself which is modeled as a linear function of predictor variables. The inverse link function is the exponential. In this model, the regression coefficient β_j represents the expected change in the log of the mean of undergraduate counts with changing household characteristics.

Since the exposure time is the same for all my subjects, a feature of the log-link allows us to express exponentiated coefficients as:

$\lambda_i = \exp(\sum X'_i\beta)$ (2)

For this model an exponentiated regression coefficient; exp $\{\beta_j\}$ represents a multiplicative effect of the j^{th} household characteristic on the mean of undergraduate counts. Change in household characteristic multiplies the mean of undergraduate counts by a factor exp $\{\beta_j\}$. This can be interpreted as the incidence-rate ratio of undergraduate counts associated with any change of household characteristics.

Since analysis focus is on expected mean counts of undergraduates per a group of households with characteristics X_i .

Then, letting $Y_{i,j,k,l,m}$ be the number of undergraduates per the m^{th} household, of $(i, j, k, l, m)^{th}$ characteristics, where i denotes family structure, j Family size, k House Head Educational level, and l Household SES. So, $Y = \sum_n Y_{i,j,k,l,m}$ is the total count of undergraduates of all the households having $(i, j, k, l, m)^{th}$ characteristics. Then if each of the observations in this group of households is a realization of an independent Poisson variate with mean λ_{ijklm}, then the group total will also be a realization of a Poisson variate with mean $n_{ijklm}\lambda_{ijklm}$

Where n_{ijklm} is the number of individuals in households with $(i, j, k, l, m)^{th}$ Characteristic.

Postulating a log-linear model for the individual households mean counts:

$$log\lambda_{ijklm} = E\left[Y_{ijklm}\right] = X'_{ijklm}\beta$$

The log of the expected value of the group totals:

$$logE[Y] = \log(n_{ijklm}\lambda_{ijklm})$$

$$= \log\left(n_{ijklm}\right) + \log(\lambda_{ijkm})$$

$$= \log\left(n_{ijklm}\right) + X'_{ijklm}\beta$$

Let u_i represent unobserved household characteristics and measurement errors on the data and let: $E\{Y_i/X_i\} = \lambda(x_i, \beta_i, \lambda_i) = \lambda_i$

Where E stands for the expectation operator, β is the k-dimensional parameter vector to be estimated and u_i is the unobserved variables and measurement errors in the data.

The general form of the log-linear regression model specification would then be:

$$\log(\lambda_{ijkm}) = \log(n_{ijklm}) + X'_{\text{ijklm}}\beta + u_i$$

Thus, the group expected counts follow a log-linear model with exactly the same coefficient β as the individual mean counts, except for the fact that the linear predictor includes the term $\log(n_{ijklm})$ referred to as the offset.

This offset takes care of the differences in the number of individuals involved in my study having respective characteristics. By including $\log(n_{ijklm}) = \varphi$ as an offset in the equation, it is differentiated from other coefficients in the regression model by being carried through as a constant and forced to have a coefficient of 1.0.

The final Poisson regression model thus estimated is:

$$G_i = \varphi e^{\beta_0+\beta_{i1}X_1+\beta_{i2}X_2+\cdots+\beta_{ij}X_j+\sigma\varepsilon_{ij}}$$

Where G_i is number of undergraduates, $\varphi_i$is the logarithm of the number of households, β is the vector of parameters for the various family structures affecting number of undergraduates, while X represents the characteristics of interest.

This final model falls within the framework of generalized linear models described by Nelder and Wedderburn (1972), representing a special case of error or stochastic structure, which is

Poisson distributed. The logarithmic link function between the expectation of the rate of undergraduate counts and household characteristics including an offset allows for the estimation of maximum likelihood, standard errors, likelihood ratio and goodness-of-fit chi-squared statistics.

3.2.1 Maximum Likelihood Estimation

Assuming independent Poisson distribution of the number of undergraduates; Y:

$$f[Y] = e^{-\lambda}\frac{\lambda^y}{y!},\ y = 0,1,2,....$$

Taking logs:

$$lnf(y) = -\exp(\lambda) + y\lambda - \ln(y!)$$

$$= -\exp(x'\beta) + yx'\beta - \ln(y!)$$

The log-likelihood is:

$$L(\beta) = \sum_{i=1}^{n}\{-\exp(x'\beta) + yx'\beta - \ln(y!)\}$$

$$\frac{\partial L(\beta)}{\partial \beta} = \sum_{i=1}^{n}\{-\exp(x'\beta)x + yx\}$$

Maximum Likelihood Estimates are solutions to:

$$\sum_{i=1}^{n}\{y - \exp(x'\beta)\}x = 0$$

This equation is non-linear in β and has no analytical solution. To solve it, an iterative method is employed known as Newton- Raphson method.

Once estimates of β are obtained, the value of maximum log-likelihood is computed and is used in AIC information criteria. The smaller AIC is the better the fit.

3.2.2 Newton-Raphson Method

When a sample is taken from my Poisson distribution, the log-likelihood is:

$$L(\lambda) = \sum_i n_i log\lambda_i - \sum_i \lambda_i$$

$$= \sum_i n_i \left(\sum_j x_{ij}\beta_j\right) - \sum_i exp\left(\sum_j x_{ij}\beta_j\right)$$

The sufficient statistic for β_j is its coefficient, $\sum_i n_i x_{ij}$ since;

$$\frac{\partial}{\partial\beta_j}\left[exp\left(\sum_j x_{ij}\beta_j\right)\right] = x_{ij}exp\left(\sum_j x_{ij}\beta_j\right) = x_{ij}\lambda_i$$

Then,

$$\lambda_j = \frac{\partial L(\beta)}{\partial\beta_j} = \sum_i n_i x_{ij} - \sum_i \lambda_i x_{ij}$$

$$h_{jk} = \frac{\partial^2 L(\beta)}{\partial\beta_j\partial\beta_k} = -\sum_i \lambda_i x_{ij} x_{ik}$$

So that,

$$\lambda_j^{(t)} = \sum_i \left(n_i - \lambda_i^{(t)}\right) x_{ij}$$

And

$$\lambda_{jk}^{(t)} = -\sum_{i} \lambda_i^{(t)} x_{ij} x_{ik}$$

The t^{th}approximation $\lambda^{(t)}$for $\hat{\lambda}$ derives from $\beta^{(t)}$ through, $\lambda^{(t)} = \exp(X\beta^{(t)})$. It generates the next value $\beta^{(t+1)}$ using,$\sum_i (y_i - \lambda_i) x_{ij} = 0$, which in this context is:

$$\beta^{(t+1)} = \beta^{(t)} + \left[X' diag(\lambda^{(t)}) X\right]^{-1} X'(n - \lambda^{(t)})$$

This in turn produces,$\lambda^{(t+1)}$, and so on.

3.2.3 Test for goodness of fit

Pearson's chi-square statistic:

$$x2 = \sum_i \left(\frac{Y_i - \lambda^{\wedge}}{\sigma_i^2}\right)^2 \sim \chi_{n-p}^2$$

Is used to determine if the inclusion of; - Family Structures, Highest level of education for household head, Number of usual members of a household (Family size) and SES index of a household have significant association with count of undergraduates in a household.

3.3 Zero-inflated Poisson Modeling

For this study, the population consists of two observations (states); the first is based on zero counts for instances where no individual in a given household was of age to have undertaken undergraduate studies as of the time the census exercise took place. The second observation is a

Poisson process of undergraduate counts for households that had individuals who were of age to have undertaken undergraduate studies as of the time census took place. These two states generate more zeros than can be predicted by the standard Poisson regression model which as a result can lead to an overall poor fit. Zero-inflated Poisson regression model (ZIP) may be more appropriate in this scenario.

The first observation comprises of a binary process that leads to zero counts known as structural zeros occurring with a probability, p_i . The second observation is a Poisson process that generated counts of zero or greater than zero with a probability $1 - p_i$. Zeros in this state are known as sampling zeros.

In consideration of these two states let Y_i be independent random number of undergraduates from households with Zero-inflated count distribution (ZIP), then model Y_i as a mixture as follows:

$$Y_i = \left\{ \begin{array}{ll} 0, & \text{with probability } p_i \\ \lambda, & \text{with probability } 1 - p_i \end{array} \right\}$$

Implying that the first state that has households with no individuals who are of age to undertake undergraduates studies occurs with probability p_i resulting into zero counts and the second state where there are households with individuals who are of age to be undergraduates will constitute the Poisson process that has expected mean count λ of undergraduates occurring with probability;

$1 - p_i$.

Then, Y_i being a non-negative zero-inflated random variable has a ZIP distribution denoted as $Y_i \sim ZIP(\lambda, p_i)$ and expressed as:

$$\Pr(Y_i = y) = \begin{matrix} p_i + (1 - p_i)\Pr(K_i = 0), y = 0 \\ (1 - p_i)\Pr(K_i = y), y => 0,\ 0 \le p_i \le 1 \end{matrix}$$

$$= \left\{ \begin{matrix} p_i + (1 - p_i)\exp(-\lambda), y = 0 \\ (1 - p_i)\frac{\exp(-\lambda)\lambda^y}{y!}, y => 0,\ 0 \le p_i \le 1 \end{matrix} \right\}$$

One of the properties of Poisson distribution is Po~(y, 0) = 0 for all y>0 and Po~ (0, 0) = 1. Therefore, Po~(y, 0) is the one point distribution putting all its mass at zero. Thus, ZIP distribution is a mixture model of point mass at zero and Poisson distribution. The first part models the structural zeros and commonly uses logistic regression and the second part models Poisson distribution conditional on excess zeros; that is the sampling zeros and actual undergraduate counts.

The implication of this distribution is that; the overall probability of having no undergraduate from a household is a combination of probabilities of zeros from each state described, weighted by the probability of being in that state which, is a Poisson chance, i.e. $p_i + (1 - p_i)\exp(-\lambda)$. On the other hand, the probability of having an undergraduate in a household is given by probability of being in the second state weighted by probability of the Poisson realization;

I.e. $(1 - p_i)\frac{\exp(-\lambda)\lambda^y}{y!}, y => 0$

Properties of the ZIP distribution are:

Mean:

$E[Y_i] = (1 - p_i)\lambda_i$

It can be seen that if p_i equal zero, ZIP model reduces to the standard Poisson model.

Variance:

$V(Y_i) = E[Y_i](1 + p_i\lambda_i)$

Implying as p_i approaches one, the variance increases and the data exhibits greater over-dispersion.

Since p_i depends on the characteristics of observed household i, p_i is written as a function of $x'_i\beta$ where x'_i is a vector of household characteristics and β is a vector of coefficient parameters to be estimated. The function that relates the product $x'_i\beta$ and probability p_i is called the zero-inflated link function and can be specified either as the logistic function or the probit function i.e.

$$\text{logit}(p(x_i)) = \alpha_o + \beta_o(x_i)$$

$$\log(\lambda(x_i)) = \alpha_1 + \beta_1(x_i)$$

Where $\lambda(x_i)$ is the mean count of undergraduates expressed as a function of household characteristics; x_i of interest through a log transformation. It is assumed that same set of household characteristic is affecting zero counts in both states. α_oAnd α_1 are unknown intercept parameters for each regression component and $\beta_o$and β_1 are vectors representing coefficients to be estimated for the various household characteristics.

Thus, the distribution of the number of undergraduates; y_i conditional on the household characteristics; x_i is modeled as:

$$(1-p_i)Po\sim(y,0) + p_iPo\sim(y,\lambda) = (1-p_i)Po\sim(y,0) + p_iPo\sim(y,\exp(\alpha_i + x'\beta))$$

Since we know;

$$E[y] = \exp(\alpha + x'\beta)$$

3.3.1 Maximum Likelihood Estimation

Taking n observations, the log-Likelihood function for undergraduate counts y_i is:

$$L(\lambda, p_i) = \sum_{i=1}^{n} (I_y = 0 \log(p_i + (1 - p_i)e^{-\lambda}) + I_y > 0 \, log(1 - p_i)\frac{e^{-\lambda}\lambda^y}{y!}))$$

$$= \sum_{i=1}^{N} (I_y = 0 \log(p_i + (1 - p_i)e^{-\lambda}) + I_y > 0 \log(1 - p_i) + y_i \log\lambda_i - \lambda_i - \log y_i!))$$

The expression I_y is the indicator function for my two states of observations. I.e. is equal to 1 if the observation is true and 0 otherwise. Parameters p_i and λ_i can be estimated using the link functions:

$\log\lambda = B\beta$ And

$$\log\left(\frac{p_i}{1 - p_i}\right) = G\gamma$$

Which Lambert (1992) suggested when it comes to applying ZIP in practical modeling situations:

B and G are matrices for household characteristics under my study. β And γ are parameters to be estimated using either Newton Raphson or Fisher scoring methods. Fisher scoring method is however preferred because the second derivative of L can be simplified by taking expectations.

3.3.2 Method of Fisher Scoring

Assuming λ and p_i are not functionally related, the first and second derivatives of $L(\lambda, p_i)$ with respect to β and p_i are:

$$\frac{\partial \ell}{\partial \beta_j} = \frac{\partial \ell}{\partial \lambda_i}\frac{\partial \lambda_i}{\partial \beta_j}$$

$$= \sum_{i=1}^{n}\{I_{(y_i=0)}\left[\frac{-(1-p_i)e^{-\lambda_i}}{p_i+(1-p_i)e^{-\lambda_i}}\right]\lambda_i + I_{(y_i>0)}(y_i - \lambda_i)\}x_{ij},$$

$j = 0,1,2, \dots p$

$$\frac{\partial \ell}{\partial p_j} = \sum_{i=1}^{n}\{I_{(y_i=0)}\left[\frac{(1-e^{-\lambda_i})}{p_i+(1-p_i)e^{-\lambda_i}}\right] + I_{(y_i>0)}[\frac{-1}{1-p_i}]\};$$

$$\frac{\partial^2 \ell}{\partial \beta_j \partial \beta_k} = \sum_{i=1}^{n}\{I_{(y_i=0)}\left[\frac{-e^{-\lambda_i}[(1-\lambda_i)p_i + (1-p_i)e^{-\lambda_i}](1-p_i)\lambda_i}{[p_i + (1-p_i)e^{-\lambda_i}]^2}\right] + I_{(y_i>0)}(-\lambda_i)\}x_{ij}x_{ik},$$

$j = 0,1,2, \dots p$

$$\frac{\partial^2 \ell}{\partial p_i^2} = \sum_{i=1}^{n}\{I_{(y_i=0)}\left[\frac{-(1-e^{-\lambda_i})^2}{[p_i+(1-p_i)e^{-\lambda_i}]^2}\right] + I_{(y_i>0)}[\frac{-1}{(1-p_i)^2}]\};$$

$$\frac{\partial^2 \ell}{\partial \beta_j \partial p_i} = \frac{\partial^2 \ell}{\partial p_i \partial \beta_j} = \sum_{i=1}^{n}\{I_{(y_i=0)}\left[\frac{\lambda_i e^{-\lambda_i}}{[p_i + (1-p_i)e^{-\lambda_i}]^2}\right]\}x_{ij}$$

Using the fact that

$E = [I_{(y_i=0)}] = P_r(Y_i = 0) = p_i + (1-p_i)e^{-\lambda_i}$ And

$E = [I_{(y_i>0)}] = P_r(Y_i > 0) = (1-p_i)(1-e^{-\lambda_i})$

Then,

$$-E = \frac{\partial^2 \ell}{\partial \beta_j \partial \beta_k} = \sum_{i=1}^{n}\{I_{(y_i=0)}\left[\frac{-e^{-\lambda_i}[(1-\lambda_i)p_i + (1-p_i)e^{-\lambda_i}](1-p_i)\lambda_i}{[p_i + (1-p_i)e^{-\lambda_i}]^2}\right]$$

$$+ I_{(y_i>0)}(-\lambda_i)\}x_{ij}x_{ik}, \qquad j = 0,1,2, \dots p$$

$$-E = \frac{\partial^2 \ell}{\partial p_i^2} = \sum_{i=1}^{n}\{I_{(y_i=0)}\left[\frac{-(1-e^{-\lambda_i})^2}{[p_i + (1-p_i)e^{-\lambda_i}]^2}\right] + I_{(y_i>0)}[\frac{-1}{(1-p_i)^2}]\}$$

$$-E = \frac{\partial^2 \ell}{\partial \beta_j \partial p_i} = \sum_{i=1}^{n} \{I_{(y_i=0)} \left[\frac{\lambda_i e^{-\lambda_i}}{[p_i + (1-p_i)e^{-\lambda_i}]^2}\right]\} x_{ij}$$

Hence the estimates of β and p_i at $(m+1)^{th}$ iteration denoted by β^{m+1} and $p_i{}^{m+1}$, are given by:

$$\begin{pmatrix} \beta^{(m+1)} \\ p_i^{(m+1)} \end{pmatrix} = \begin{pmatrix} \beta^{(m)} \\ p_i^{(m)} \end{pmatrix} + \left[\mathcal{T}^{(m)}(\beta, p_i)\right]^{-1} s^{(m)}(\beta, p_i),$$

Where the score vector and the expected information matrix respectively, evaluated at $\beta = \beta^{(m)}$ and $p_i = p_i^{(m)}$ are as follows:

$$s(\beta, p_i) = \begin{pmatrix} s_\beta(\beta, p_i) \\ s_{p_i}(\beta, p_i) \end{pmatrix} = \begin{pmatrix} \frac{\partial \ell}{\partial \beta} \\ \frac{\partial \ell}{\partial p_i} \end{pmatrix},$$

$$\mathcal{T}(\beta, p_i) = \begin{vmatrix} \mathcal{T}_{\beta\beta}(\beta, p_i) & \mathcal{T}_{\beta p_i}(\beta, p_i) \\ \mathcal{T}_{p_i\beta}(\beta, p_i) & \mathcal{T}_{p_i p_i}(\beta, p_i) \end{vmatrix}$$

Where the elements $\mathcal{T}_{\beta\beta}, \mathcal{T}_{\beta p_i} = \mathcal{T}_{p_i\beta}$ and $\mathcal{T}_{p_i p_i}$ are, respectively,

$$-E\left[\frac{\partial^2 \ell}{\partial \beta \partial \beta^T}\right], \; -E\left[\frac{\partial^2 \ell}{\partial \beta \partial p_i}\right], \text{ and } -E\left[\frac{\partial^2 \ell}{\partial {p_i}^2}\right].$$

With good starting values $\beta^{(0)}$, $p_i{}^{(0)}$ and hence $\lambda^{(0)}$, $p_i{}^{(0)}$ the iterative scheme converges in a few step, convergence is obtained with a stopping rule, such as, $\left|\ell^{(m+1)} - \ell^{(m)}\right| \le \epsilon$, where $\ell^{(m)}$ and $\ell^{(m+1)}$ are the log-likelihood, $\ell(\lambda, \beta; y)$ evaluated using the estimates of λ and p_i from the (m) and (m+1) iterations, respectively. The asymptotic variance-covariance matrix for $(\hat{\beta}. \hat{p}_i)$ is automatically provided in the final iteration.

4 Chapter 4: DATA ANALYSIS AND RESULTS

4.1 Data

2009 Kenya National Housing and Population Census offer valuable data that can be used to evaluate the effect of family structures on education attainment. During the census the country was divided into small counting units called Enumeration Areas (EAs) comprising of an average of 100 households, by cartographic mapping for purposes of enumerating all people within Kenyan boundaries. The EAs were then categorized into four; - (a) EA in settled agricultural areas, (b) Urban/Peri-Urban (c) Arid and Semi-Arid areas (d) Forests/National parks or Game Reserves. These EA categories were further broken down to EA Type and EA Status. Households were also categorized into conventional and non-conventional.

This study focuses on conventional households from EA Type 3 (Peri-Urban) and EA Status 9 (Formal Settlement) only to avoid the problem of clumping and truncation when it comes to variables selection for construction of Social Economic Status (SES) index of individual households. Confined to only semi-urban areas of the country is in a way also trying to bring some homogeneity in the quality of schools which is not available as a variable from the census data for my study.

Social economic status of a household as a variable is also not available from the census data but various variables are available from the census data that can be used in its construction which include;- highest level of education completed by the household head, main economic activity for the household head, Tenure status, Main material for roofing, Main material for wall, Main material for floor, Main waste disposal method, Main source of water, Main source of fuel, Main source of lighting, and ownership of Radio, TV, Mobile Phone, Landline Telephone, Computer, Bicycle, Motor Cycle, Car, Truck Lorry, Tractor, Bus, Refrigerator, Boat, Animal Drawn cart, Canoes, Tuk Tuk, Exotic cattle, Indigenous cattle, Sheep, Goat, Camel, Donkey, Pig, Indigenous chicken, Commercial chicken Bee hives, and other livestock. Education attainment is assumed to be a function of a set independent household and demographic variables and therefore other variables of importance from the census include; - Marital of household head, Size of usual members of a household and Sex of household head.

Analysis is at the household level where the family structure of a particular household is indicated by the marital status of the household head. The numbers of children who were indicated to be usual members of the household and had completed their undergraduate studies are regressed against the various family structures controlling and not controlling for the various background household and demographic variables deemed to also influence education attainment of a child in a particular household.

4.1.1 Dependent measure: Education attainment measure

The response variable for this study is the number of usual members of each household who have completed their undergraduate studies as indicated during the census.

4.1.2 Independent measures

The primary independent variables are family structures and sex for household head of the household. **Family structure** is measured by a set of categorical variables of marital status of household head in the census data. This categories are;- 1= Never married, 2= Married monogamous, 3= Married polygamous, 4= Widowed, 5= Divorced and 6= Separated. The primary interest of this study is to look at effect of family structures on education attainment of children in a household. Along with looking at effect of family structures, distinction is made between; Effect with regard to **Sex** for household head which, is a dichotomous variable in the census data where Male = 1 and Female = 2. Female-headed household and Male-headed household family structures are likely to affect children differently.

4.1.3 Other Background household variables

Several control variables are included in the multivariate analysis because previous studies have found them to be associated with education attainment. These include; - SES index of a household, Number of usual members of a household and Highest educational level completed by the household head.

SES index of household is constructed using available variables in the census data. A better SES index for a household means more resources at the disposal of a child in the household and thus a better education attainment and the converse is also true. For this study, households have been categorized into; - Poor, Middle class and Rich depending on the constructed SES Index. The study deliberately avoided variables on the number of dwelling units and habitable rooms to avoid them being unjustifiably weighted more during analysis because it is possible to have fewer households having mud-walled houses with many habitable rooms. Such a scenario may lead to the variable being weighted more than the many stone-walled houses having say; two to three habitable rooms. Main employer variable for household head was also avoided because by including, it would have meant I am dealing with only households which had household heads working for pay.

When it comes to the main activity of household head, similar variables like, worked for pay; 1, on leave; 2, and Sick leave; 3 are combined together as those working for pay. Seeking working work; 8, seeking work no action; 9, No work available; 10, Homemaker; 12, Full time student; 13, Incapacitated; 14 will also be combined together as not employed.

Highest educational level completed by a household head and usual members of the household is categorical data in the census data coded 97, 96, 0, 1 through to 26. There are several possible reasons why children raised by parents with higher education levels attain higher levels of education themselves than children with less educated parents. One theoretical explanation for the relationship between parents' schooling and their children's educational success suggests that parents invest in their children by providing them both economic resources and human capital (Becker, 1981). This study has re-coded the various categories for household head education completion into five categories;- Illiterate (those who have never been to school;97, those Not completed basic/post literacy;21, those attending Madrass/Duksis;25, those that have completed pre-primary;96, have completed basic/post literacy;22, completed Madrassa/Duksis;26, and those Not completed standard one;0), Primary school level (those Not completed youth polytechnic;23, have completed standard one to form one;1 - 9), Secondary school level (Not completed post secondary;15, those indicated to have completed youth polytechnic;24, and Completed, Form two to four;10 - 12), College level (those indicated Not completed undergraduate;17,those who have completed Form five;13, Form six;14 and have completed post secondary;16), University level (those who have completed undergraduate;18, Attending/Completed masters;19, Attending/Completed PHD;20)

Usual members of a household is a dichotomous variable of 1= Yes and 2= No. For each household, focus will be on those who have completed undergraduate studies and are usual members of household, at the same time looking at the total number of all those who are usual members in a household. Previous research has shown the inverse relationship between family size and education achievement is due to the dilution of available family resources in larger families compared to the resulting concentration of resources in smaller families; Alwin and Thornton (1984), Blake (1989), Zajonc (1976).

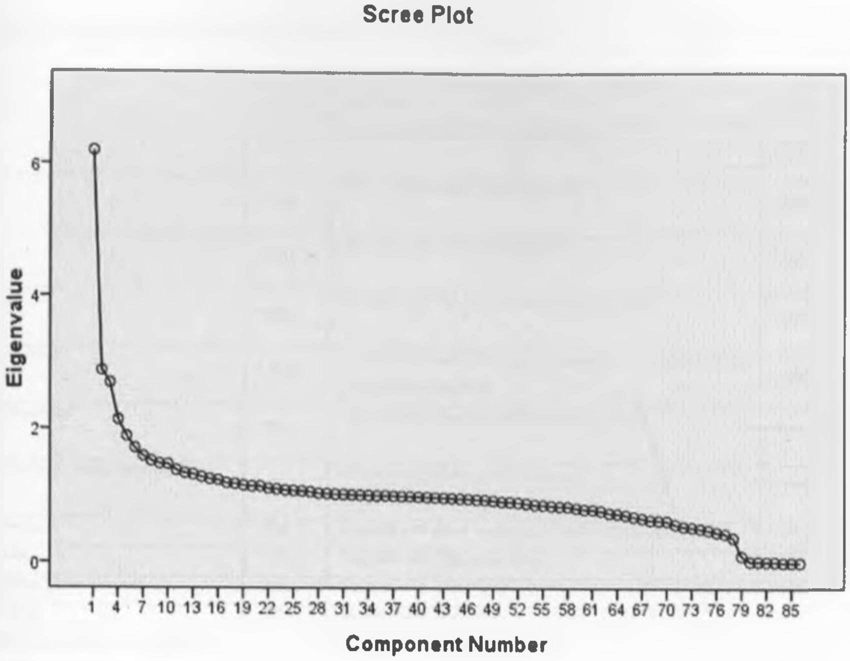
4.1.4 Missing Data

Missing information during census for categorical data was coded "9" and for continuous data was coded "99" or "9999".

4.2 Constructed SES for households

In constructing a Social Economic Status variable for the households, a scree plot has been obtained for the selected variables as presented in Figure: 4.1, below.

Figure 4.1



From the above plot, it is evident that the first 5 principle components account for most of the variability in SES among households.

After principle component analysis of all selected variables, it was found that the first principle component explains the Basic aspect of social economic status of a household while sophisticated aspect of social economic status of a household was better explained by the second principle component. I.e. sophisticated assets like a car, refrigerator, Tuk tuk, TV, Radio, Mobile phone, Landline, bicycle, Tuck/Lorry, that cannot be classified as basic needs in life were scored negatively in the first principle component but positively in the second principle component. As such they were excluded from my first Principle Component analysis.

Table: 4.1 below show the first Principle Component scoring of basic SES indicator variables for the households while leaving out the sophisticated SES indicators variables.

Table: 4.1 Selected Household Variables First PCA Scores:

Variable	Score	Variable	Score	Variable	Score
H17 Sheep	- 022	Houses with Main Sewer disposal	252	Houses with Electricity as main source of lighting	647
H17 Goat	- 058	Houses with Septic Tank disposal	306	Houses with Lantern as main source of lighting	.011
H17 Donkeys	- 001	Houses with Cess Pool disposal	.156	Houses with Tin Lamp as main source of lighting	-.473
H17 Pigs	007	Houses with VIP Pit Latrine waste disposal	269	Houses with Firewood Lighting as main source of lighting	-.063
H17 Indigenous chicken	- 020	Households with Pit latrine covered or uncovered as main waste disposal	- 208	Houses with Solar Lighting as main source of lighting	080
H17 Chicken Commercial	039	Houses with Bucket Latrine waste disposal	025	Houses with Other sources as main source of lighting	027
Households with exotic or indigenous cattle	- 049	Houses with Bush waste disposal	- 232	Households that use pressure or gas lamps as main source of lighting	.000
H17 Other Livestock	002	Houses with other methods of waste disposal	- 015	Government's rented House	205
Tiles roofed Houses	150	Houses with Cemented floor	762	Parastatal's rented House	.108
Houses with Concrete slabs as roofs	139	Houses with Tilled floor	158	Local authority's rented House	030
Houses with Tin roofs	- 021	Houses with wood floor	005	Private company's rented House	264
Households with corrugated iron sheets or asbestos as main roofing material	191	Houses with Earth floor	- 785	Individual's rented House	422
ther types of roofs	- 051	Other types of floors	- 015	NGO's rented House	024
n sheets walled Houses	.146	Houses with Electricity as main source of cooking fuel	145	Other type of tenure	.027
use with Tin walls	- 007	Houses with LPG as main source of cooking fuel	297	Stone walled Houses	.508
r types of walls	- 055	Houses with Firewood as main source of cooking fuel	- 732	Brick or block walled Houses	248
r from Borehole	045	Houses with Other sources as main source of cooking fuel	046	Houses with Mud or Wood walls	- 590
n Water harvesting	080	Households that use charcoal or paraffin as main source of fuel	633	Houses with Mud mixed with Cement walls	- 037
r from Vendors	170	Household head that is illiterate	- 223	Wood walled Houses	044
holds with piped water into dwelling or in compound	431	Household head whose highest level of education is primary level	-.276	Housed head who draws a salary	388
hold head whose highest level of ucation is university level	227	Household head whose highest level of education is secondary level	302	Household head that is unemployed	-.082
hold head who farms	-.378	Household head whose highest level of education is college level	287	Household head that is in business	070
hold head that is retired	044				

Table: 4.2 below Shows comparison scoring of all initially selected household variables by the first and second Principle Components.

Table: 4.2 Selected Household variables First Two PCA Scores:

riable	PCA 1 Score	PCA 2 Score	Variable	PCA 1 Score	PCA 2 Score
17. Sheep	- 007	-.069	Other type of tenure	.018	.009
17. Goat	-.036	-.137	Tiles roofed Houses	180	-.238
17. Donkeys	-3.365E-5	-.003	Houses with Concrete slabs as roofs	.133	-.063
17. Pigs	.008	-.003	Houses with Tin roofs	-.021	-.039
17. Indigenous chicken	.018	-.054	Other types of roofs	-.057	-.138
17. Chicken Commercial	053	-.050	Stone walled Houses	491	.010
17. Other	006	-.012	Brick or block walled Houses	236	127
28. Radio	-.256	-.094	Houses with Mud or Wood walls	-.554	-.187
28. TV	-.574	.041	Houses with Mud mixed with Cement walls	-.035	.021
28. Mobile Phone	-.461	-.112	Wood walled Houses	.040	.115
Landline Telephone	-.189	285	Iron sheets walled Houses	.121	137
Computer	- 354	.364	House with Tin walls	- 012	-.016
Bicycle	-.130	.053	Other types of walls	-.058	-.108
Motor Cycle	-.152	.129	Houses with Cemented floor	.727	.271
Car	- 394	416	Houses with Tilled floor	.204	-.242
Truck/Lorry/Tractor/Bus	-.189	.395	Houses with wood floor	.007	-.016
Refrigerator	- 425	513	Houses with Earth floor	- 758	-.217
Boat	-.114	489	Other types of floors	-.018	-.057
Animal Drawn Cart	-.106	.200	Water from Borehole	.052	.021
Canoes	-.103	494	Rain Water harvesting	.099	-.037
Tuk tuk	-.116	484	Water from Vendors	.136	.123
mment's rented House	.197	-.067	Houses with Main Sewer disposal	.245	-.104
ial's rented House	.104	-.038	Houses with Septic Tank disposal	335	-.229
authority's rented House	.027	.002	Houses with Cess Pool disposal	134	.012
company's rented House	.219	.062	Houses with VIP Pit Latrine waste disposal	.270	-.049
al's rented House.	.329	.266	Houses with Bucket Latrine waste disposal	.019	-.060
rented House	.024	-.105	Houses with Bush waste disposal	-.237	-.313
with Electricity as main source of cooking fuel	.159	-.108	Houses with other methods of waste disposal	-.019	-.023
with LPG as main source of cooking fuel	340	-.248	Houses with Lantern as main source of lighting	.006	.264
with Firewood as main source of cooking fuel	- 664	-.182	Houses with Tin Lamp as main source of lighting	- 487	-.177
with Other sources as main source of cooking	.036	.020	Houses with Firewood Lighting as main source of lighting	-.072	-.129
with Electricity as main source of lighting	.663	-.069	Houses with Solar Lighting as main source of lighting	.118	-.011
with Other sources as main source of lighting	.018	-.004	Household head whose highest level of education is secondary level	.298	.138

Table: 4.2 Continued.........

riable	PCA 1 Score	PCA 2 Score	Variable	PCA 1 Score	PCA 2 Score
ousehold head whose highest level of education university level	.281	-.220	Households with piped water into dwelling or in the compound.	.409	.047
ousehold head who is illiterate	-.248	-.217	Household head whose highest level of education is college level	.319	-.063
oused head who draws a salary	.351	.122	Household head who is retired	.056	-.026
ousehold head who is unemployed	-.104	-.049	Households that use charcoal or paraffin as main source of fuel.	.544	.298
ousehold head who is in business	.078	.036	Households that use pressure or gas lamps as main source of lighting.	-.002	-.046
ousehold head who farms	-.334	-.099	Households with Pit latrine covered or uncovered as main waste disposal	-.206	.361
ouseholds with corrugated ironsheets or sbestos as main roofing material	.185	.366	Households with exotic or indigenous cattle.	.000	-.121

It clearly shows the two different aspects of SES for a household that the two components are explaining when all the variables are included.

Table: 4.3 shows a total variation of 11.28% in SES among households accounted for by the first two components.

Table: 4.3

Component	Initial Eigen values			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	6.101	7.533	7.533	6.101	7.533	7.533
2	3.037	3.750	11.282	3.037	3.750	11.282

This is relatively low and could be as a result of other unobserved variables in a household that are key in determining SES of a household. A variable like religion that was not available in my data is likely to be key determinant of SES of a household in the Kenyan scenario.

4.3 Poisson Regression Fitted Models

Ascertaining if really indeed differences exists between the number of undergraduates from the different Family Structures a Main effect Model with number of undergraduates as the response variable and Family Structures as the factors was fitted and Table: 4.4 below shows pair wise comparison of the expected mean of undergraduates from the different Family Structures.

Table: 4 4 Pair wise Comparisons

(I) Family structure	(J) Family structure	Mean Difference (I-J)	Std Error	df	Sig	95% Wald Confidence Interval for Difference	
						Lower	Upper
Never Married	Married Monogamous	.00[a]	.000	1	.017	.00	.00
	Married Polygamous	.00[a]	.000	1	.000	.00	.00
	Widowed	.00[a]	.000	1	.000	.00	.00
	Divorced	.00[a]	.000	1	.000	.00	.00
	Separated	.00[a]	.000	1	.000	.00	.00
Married Monogamous	Never Married	.00[a]	.000	1	.017	.00	.00
	Married Polygamous	.00[a]	.000	1	.000	.00	.00
	Widowed	.00[a]	.000	1	.000	.00	.00
	Divorced	.00[a]	.000	1	.006	.00	.00
	Separated	.00[a]	.000	1	.000	.00	.00
Married Polygamous	Never Married	.00[a]	.000	1	.000	.00	.00
	Married Monogamous	.00[a]	.000	1	.000	.00	.00
	Widowed	.00[a]	.000	1	.006	.00	.00
	Divorced	.00	.000	1	.675	.00	.00
	Separated	.00[a]	.000	1	.003	.00	.00
Widowed	Never Married	.00[a]	.000	1	.000	.00	.00
	Married Monogamous	.00[a]	.000	1	.000	.00	.00
	Married Polygamous	.00[a]	.000	1	.006	.00	.00
	Divorced	.00	.000	1	.470	.00	.00
	Separated	.00[a]	.000	1	.000	.00	.00
Divorced	Never Married	.00[a]	.000	1	.000	.00	.00
	Married Monogamous	.00[a]	.000	1	.006	.00	.00
	Married Polygamous	.00	.000	1	.675	.00	.00
	Widowed	.00	.000	1	.470	.00	.00
	Separated	.00[a]	.000	1	.039	.00	.00
Separated	Never Married	.00[a]	.000	1	.000	.00	.00
	Married Monogamous	.00[a]	.000	1	.000	.00	.00
	Married Polygamous	.00[a]	.000	1	.003	.00	.00
	Widowed	.00[a]	.000	1	.000	.00	.00
	Divorced	.00[a]	.000	1	.039	.00	.00

The results show significant differences between most of the family structures except for pairs like; widowed & divorced and divorced & polygamous. The implication of this is that effects on education attainment for “Widowed” & “Divorced” and “Divorced” & “Polygamous” Family is the same.

Table: 4.5 below show goodness of fit results for the Main effect Model that tests the direct effects of Family Structures on counts of undergraduates.

Table: 4.5

Goodness of fit for Main effects Model of Family Structures			
	Value	DF	Value/DF
Deviance	4.350E4	628793	.069
Scaled Deviance	4.350E4	628793	
Pearson Chi-Square	6.624E5	628793	1.054
Scaled Pearson Chi-Square	6.624E5	628793	
Log Likelihood[a]	-2.596E4		
Akaike's Information Criterion (AIC)	5.194E4		
Finite Sample Corrected AIC (AICC)	5.194E4		
Bayesian Information Criterion (BIC)	5.201E4		
Consistent AIC (CAIC)	5.201E4		

Test of Model Effects			
Source	Type III		
	Wald Chi-Square	DF	Sig.
(Intercept)	16860.030	1	.000
Family structure	123.948	5	.000

Omnibus Test		
Likelihood Ratio Chi-Square	DF	Sig.
143.230	5	.000

The Pearson’s chi-squares value/df is 1.054 which is an indication that Poisson assumptions are meet in model fitting.

Table: 4.6 below shows goodness of fit results for the Model that tests effects of Family Structures while considering Family size.

Table: 4.6

Goodness of fit for Main effects Model of Family Structures considering Family Size			
	Value	DF	Value/DF
Deviance	4.347E4	628792	.069
Scaled Deviance	4.347E4	628792	
Pearson Chi-Square	6.546E5	628792	1.041
Scaled Pearson Chi-Square	6.546E5	628792	
Log Likelihood[a]	-2.595E4		
Akaike's Information Criterion (AIC)	5.191E4		
Finite Sample Corrected AIC (AICC)	5.191E4		
Bayesian Information Criterion (BIC)	5.199E4		
Consistent AIC (CAIC)	5.200E4		
Test of Model Effects			
Source	Type III		
	Wald Chi-Square	DF	Sig.
(Intercept)	15083.3	1	.000
Family structure	122.280	5	.000
Family Size	12.613	1	.000
Omnibus Test			
Likelihood Ratio Chi-Square		DF	Sig.
172.875		6	.000

The Pearson's chi-squares value/df is 1.041 which is an indication that Poisson assumptions are meet in model fitting

Table: 4.7 below shows goodness of fit results for the Model that tests effects of Family Structures while considering the SES of a household.

Table: 4.7

<table>
<tr><th colspan="5">Goodness of fit for Main effects Model of Family Structures considering Social Economic status of a household</th></tr>
<tr><td colspan="2"></td><td>Value</td><td>DF</td><td>Value/DF</td></tr>
<tr><td colspan="2">Deviance</td><td>3.747E4</td><td>625969</td><td>.060</td></tr>
<tr><td colspan="2">Scaled Deviance</td><td>3.747E4</td><td>625969</td><td></td></tr>
<tr><td colspan="2">Pearson Chi-Square</td><td>5.956E5</td><td>625969</td><td>.951</td></tr>
<tr><td colspan="2">Scaled Pearson Chi-Square</td><td>5.956E5</td><td>625969</td><td></td></tr>
<tr><td colspan="2">Log Likelihood[a]</td><td>-2.292E4</td><td></td><td></td></tr>
<tr><td colspan="2">Akaike's Information Criterion (AIC)</td><td>4.586E4</td><td></td><td></td></tr>
<tr><td colspan="2">Finite Sample Corrected AIC (AICC)</td><td>4.586E4</td><td></td><td></td></tr>
<tr><td colspan="2">Bayesian Information Criterion (BIC)</td><td>4.595E4</td><td></td><td></td></tr>
<tr><td colspan="2">Consistent AIC (CAIC)</td><td>4.596E4</td><td></td><td></td></tr>
<tr><th colspan="5">Test of Model Effects</th></tr>
<tr><td rowspan="2">Source</td><td colspan="4">Type III</td></tr>
<tr><td>Wald Chi-Square</td><td>DF</td><td colspan="2">Sig.</td></tr>
<tr><td>(Intercept)</td><td>11855.735</td><td>1</td><td colspan="2">.000</td></tr>
<tr><td>Family structure</td><td>91.994</td><td>5</td><td colspan="2">.000</td></tr>
<tr><td>Social Economic Status</td><td>6147.945</td><td>2</td><td colspan="2">.000</td></tr>
<tr><th colspan="5">Omnibus Test</th></tr>
<tr><td colspan="2">Likelihood Ratio Chi-Square</td><td>DF</td><td colspan="2">Sig.</td></tr>
<tr><td colspan="2">5931.285</td><td>7</td><td colspan="2">.000</td></tr>
</table>

The Pearson's chi-squares value/df is 0.951 which is an indication that Poisson assumptions are meet in model fitting

Table: 4.8 below shows goodness of fit results for the Model that tests effects of Family Structures while considering the Highest Educational level of household head.

Table: 4.8

Goodness of fit for Main effects Model of Family Structures considering the highest level of educational attainment for the household head			
	Value	DF	Value/DF
Deviance	3.228E4	626989	.051
Scaled Deviance	3.228E4	626989	
Pearson Chi-Square	5.937E5	626989	.947
Scaled Pearson Chi-Square	5.937E5	626989	
Log Likelihood[a]	-2.034E4		
Akaike's Information Criterion (AIC)	4.069E4		
Finite Sample Corrected AIC (AICC)	4.069E4		
Bayesian Information Criterion (BIC)	4.082E4		
Consistent AIC (CAIC)	4.083E4		

Test of Model Effects			
Source	Type III		
	Wald Chi-Square	DF	Sig.
(Intercept)	9807.604	1	.000
Family structure	77.538	5	.000
Educational Level	13903.373	5	.000

Omnibus Test		
Likelihood Ratio Chi-Square	DF	Sig.
11196.636	10	.000

The Pearson's chi-squares value/df is 0.947 which is an indication that Poisson assumptions are meet in model fitting.

Table: 4.9 below shows goodness of fit results for the Full Model that tests effects of Family Structures while considering all the background household variables.

Table: 4.9

Goodness of fit for Full Model of Family Structures controlling for family size, SES, and highest educational level for the head

	Value	DF	Value/DF
Deviance	3.139E4	625963	.050
Scaled Deviance	3.139E4	625963	
Pearson Chi-Square	5.671E5	625963	.906
Scaled Pearson Chi-Square	5.671E5	625963	
Log Likelihood[a]	-1.988E4		
Akaike's Information Criterion (AIC)	3.979E4		
Finite Sample Corrected AIC (AICC)	3.979E4		
Bayesian Information Criterion (BIC)	3.995E4		
Consistent AIC (CAIC)	3.997E4		

Test of Model Effects

Source	Type III		
	Wald Chi-Square	DF	Sig.
(Intercept)	7577.075	1	.000
Family structure	91.032	5	.000
Educational Level	6577.074	5	.000
SES	766.326	2	.000
Household Size	3.174	1	.075

Omnibus Test

Likelihood Ratio Chi-Square	DF	Sig.
12009.801	13	.000

The Pearson's chi-squares value/df is 0.906 which is an indication that Poisson assumptions are meet in model fitting

Table: 4.10 below shows parameter estimates for main effects Model of family structures on count of undergraduates.

Table 4.10

Family Structures Main effects Model

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	-7.402	.2182	-7.829	-6.974	1150.453	1	.000	.001	.000	.001
[Never married]	1.226	.2279	.780	1.673	28.953	1	.000	3.409	2.181	5.329
[Married monogamous]	1.051	.2189	.622	1.480	23.046	1	.000	2.860	1.862	4.391
[Married polygamous]	.536	.2265	.092	.980	5.599	1	.018	1.709	1.096	2.664
[Widowed]	.760	.2250	.319	1.201	11.413	1	.001	2.138	1.376	3.323
[Divorced]	.623	.2910	.053	1.194	4.592	1	.032	1.865	1.055	3.299
[Separated]	0[a]	.	.	.	.	.	.	1	.	.
(Scale)	1[b]									

The Main effect Model has all the parameter estimates significant with "Never Married" Family Structure contributing the highest increase of 1.226 in the mean log counts of undergraduates compared to "Separated" Family Structure. This can be attributed to the current trend in Kenya where educated career women are opting to remain single and independent because they are capable of offering their children good quality education and cater for their other needs all by themselves. "Married monogamous" follows with an increase of 1.051 in the mean log count of undergraduates compared to "Separated" Family Structure. "Married monogamous" Family Structure has been known from literature to be the most ideal for the children's' education attainment but going by this findings it seems like this is changing in Kenya and that's why it follows closely in also registering a high expected increase in the mean log count of undergraduates compared to "Separated" Family Structure. "Married polygamous" followed by "Divorced" Family structures have the lowest expected increase on mean log count of undergraduates compared to "Separated". The fact that all the coefficients are positive means that all the other Family Structures are better than "Separated" Family Structure in increasing the expected mean log count of undergraduates.

The Table: 4.11 below shows the main effects of Family structures while considering Family size.

Table: 4.11

Main effects model for Family Structures considering Family Size

...meter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
...rcept)	-7.356	.2186	-7.784	-6.927	1132.409	1	.000	.001	.000	.001
...r married]	1.233	.2279	.786	1.680	29.265	1	.000	3.431	2.195	5.364
...rried monogamous]	1.075	.2189	.646	1.504	24.101	1	.000	2.930	1.907	4.500
...rried polygamous]	.567	.2266	.123	1.011	6.265	1	.012	1.763	1.131	2.749
...owed]	.780	.2250	.339	1.221	12.026	1	.001	2.182	1.404	3.391
...rced]	.632	.2910	.062	1.202	4.718	1	.030	1.881	1.064	3.328
...arated]	0^a	.	.	.	.	.	.	1	.	.
...sehold Size	-.012	.0033	-.018	-.005	12.613	1	.000	.989	.982	.995
...)	1^b									

The expected mean log counts of undergraduates increases for all the Family structures in this Model compared to main effects Model. This implies that if all the Family Structures are ridden off the income burden that comes along with family size then they would all register an increase in the expected mean log count of undergraduates. The "Married polygamous" and "Divorced" Family Structures have again the least expected increase on mean log count of undergraduates compared to "Separated". The results also indicates that a decrease of -.012 in the mean log counts of undergraduates is expected for every unit increase in family size.

The Table: 4.12 shows the main effects of Family Structures while considering the highest Educational Levels of household heads.

Table 4.12

ain effects model for Family Structures considering Household head Educational Level Completed

ameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
rcept)	-3.785	.2194	-4.215	-3.355	297.608	1	.000	.023	.015	.035
er married]	.826	.2280	.379	1.273	13.124	1	.000	2.284	1.461	3.571
ried monogamous]	.607	.2189	.178	1.036	7.686	1	.006	1.835	1.195	2.818
ried polygamous]	.566	.2266	.122	1.010	6.244	1	.012	1.762	1.130	2.747
owed]	1.060	.2254	.618	1.502	22.126	1	.000	2.887	1.856	4.491
rced]	.730	.2910	.160	1.301	6.296	1	.012	2.076	1.173	3.672
rated]	0^a									
ain]	-4.317	.0678	-4.450	-4.184	4055.039	1	.000	.013	.012	.015
ary]	-4.562	.0456	-4.652	-4.473	10005.212	1	.000	.010	.010	.011
dary]	-3.404	.0423	-3.487	-3.322	6490.367	1	.000	.033	.031	.036
ge]	-1.969	.0408	-2.049	-1.889	2327.765	1	.000	.140	.129	.151
ily]	0^a	.	.	.	.	.	.	1	.	.
[illegible]	1^b									

The expected mean log count for undergraduate compared to "Separated" decreases for "Never Married" Family Structure from 1.226 in the Main effect Model to 0.826 in this Model and increases for "Widowed" Family Structure from 0.760 in the Main Model to 1.060 in this Model. This implies that stripped of the educational advantage "Never Married" Family structures had been presumed to have the mean log counts for undergraduates reduces drastically. "Widowed" Family Structure seems to be showing some resilience in mitigating against its effect on mean log count of undergraduates without the educational advantage. Expected increase in the mean log of counts for "Married Polygamous" Family Structure has remained almost the same but for "Divorced" has increased from .623 in the Main effect Model to .730 in this Model but the two Family Structures still have the least. To recall from my earlier findings, increases due to the effect of these two Family structures on mean log count of undergraduates was not significant and therefore from these results we can say that whenever these two Family Structures have educated heads, they are in a way able to mitigate against their effect on education attainment for their children. From this same Model illiterate and Primary school level heads of households are expected to contribute to a decrease in the expected mean log count of undergraduates by -4.317 and -4.562 respectively compared to University level heads holding family structures constant. Mean log of counts for undergraduates is expected to decrease by -1.969 for household heads who have College level of education compared to households with University level of education. All the parameter estimates are significant

The Table: 4.13 below show the main effects of Family structures while considering SES of a household.

Table: 4.13

Main effects model for Family Structures considering Social Economic Status (SES) of a household

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	-5.882	.2239	-6.321	-5.444	690.366	1	.000	.003	.002	.004
[Never married]	.857	.2332	.400	1.314	13.501	1	.000	2.356	1.492	3.721
[Married monogamous]	1.277	.2243	.838	1.717	32.439	1	.000	3.587	2.311	5.567
[Married polygamous]	1.167	.2319	.713	1.622	25.334	1	.000	3.213	2.040	5.063
[Widowed]	1.487	.2306	1.036	1.939	41.618	1	.000	4.426	2.817	6.954
[Divorced]	.869	.2950	.291	1.448	8.685	1	.003	2.386	1.338	4.253
[Separated]	0[a]	.	.	.	.	.	.	1	.	.
[Poor]	-4.020	.1227	-4.261	-3.780	1072.949	1	.000	.018	.014	.023
[Middle class]	-2.321	.0310	-2.382	-2.260	5613.898	1	.000	.098	.092	.104
[Rich]	0[a]	.	.	.	.	.	.	1	.	.
(Scale)	1[b]									

"Never married" again registers a further decrease from 1.226 in the Main effect Model to .857 in this Model in the expected increase on the mean log counts for undergraduates for "Never Married" Family structure compared to "Separated" Family Structure. All the other Family structure register an increase of expected mean log count of undergraduates compared to "Separated". This implies that the "Never Married" Family Structures without due advantage of income insinuated earlier would register a decrease in the expected increase due to its effect on the mean log counts for undergraduates compared to "Separated". "Married polygamous" and "Divorced" Family Structures remain to have the least increase mean log count of undergraduates compared to "Separated" Family Structure. From Model results, poor families are expected to lead to -4.020, decrease in the expected mean log counts of undergraduates compared to rich families.

The Table: 4.14 below shows the main effects of Family Structures on education attainment considering all the background household variables of Family size, SES, and Educational Level for head.

Table 4.14

Full Model of Family Structures considering Family Size, Head Educational Level and Social Economic Status of a household

rameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
ercept)	-3.662	.2250	-4.103	-3.221	264.978	1	.000	.026	.017	.040
ever married]	.740	.2333	.283	1.197	10.065	1	.002	2.096	1.327	3.311
arried monogamous]	.753	.2244	.313	1.193	11.269	1	.001	2.124	1.368	3.297
arried polygamous]	.835	.2320	.381	1.290	12.967	1	.000	2.306	1.463	3.633
lowed]	1.276	.2307	.824	1.728	30.602	1	.000	3.583	2.280	5.632
orced]	.824	.2951	.245	1.402	7.788	1	.005	2.279	1.278	4.063
parated]	0[a]									
erate]	-3.476	.0739	-3.621	-3.331	2212.283	1	.000	.031	.027	.036
ary]	-3.778	.0534	-3.882	-3.673	5009.424	1	.000	.023	.021	.025
ondary]	-2.930	.0460	-3.021	-2.840	4065.842	1	.000	.053	.049	.058
ege]	-1.686	.0419	-1.769	-1.604	1617.069	1	.000	.185	.171	.201
ersity]	0[a]									
r]	-2.122	.1297	-2.377	-1.868	267.832	1	.000	.120	.093	.154
e class]	-.966	.0373	-1.040	-.893	670.292	1	.000	.380	.354	.409
	0[a]							1		
old Size	-.004	.0021	-.008	.000	3.256	1	.071	.996	.992	1.000
	1[b]									

The expected increase on the mean log count of undergraduates, reduces for "Never Married" & "Married Monogamous" than earlier figures comparing with "Separated" Family Structure. "Widowed" expected increase on mean log count of undergraduates compared to "Separated" has increased from 0.760 in the Main effect Model to 1.276 in this full Model. The complementary effect that both parents have for each other in a "Married monogamous" Family Structure is expected to offer some income level and educational advantage just like earlier insinuated about "Never married", Family Structures compared to other Family Structures. As such, when these advantages are controlled the increase due to the effect of these two family structures on the mean log count of undergraduates compared to "Separated" is likely to be lower. "Widowed" Family Structure has again shown some resilience in mitigating against its effect on mean log count of undergraduates in the absence of income or educational advantages. Taking into account all these background household variables again Family size turns not significant in predicting counts of undergraduates.

The Table: 4.15 below show main effect Model results of Family Structures effects on mean log counts of undergraduates for Female headed households only.

Table: 4.15

Main effects model for Family Structures for Female Headed households

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
Intercept)	-1.041	.0075	-1.055	-1.026	19345.728	1	.000	.353	.348	.358
Never married]	.225	.0086	.208	.242	687.612	1	.000	1.253	1.232	1.274
Married monogamous]	-.235	.0078	-.250	-.220	915.350	1	.000	.791	.779	.803
Married polygamous]	-.213	.0084	-.229	-.196	635.727	1	.000	.808	.795	.822
dowed]	-.076	.0079	-.091	-.060	91.450	1	.000	.927	.913	.942
vorced]	-.041	.0115	-.063	-.018	12.574	1	.000	.960	.939	.982
parated]	0[a]	.	.	.	.	.	.	1	.	.
Scale)	.391[b]	.0012	.389	.394						

It is now evident that my earlier presumption that "Never Married" Family Structure was having most increase on mean log counts of undergraduates compared to "Separated" Family Structure because of the recent trend in Kenya where career women have chosen to be independent, holds ground. From this model only the "Never Married" Family Structure has a positive increase on the mean log count of undergraduates with a parameter estimate of .225. All the other Family Structures are having negative parameters because of the fact that all the females in these other Family Structures are not Heads of the households by choice engineered by realization they can manage their own affairs unlike the case for heads in "Never Married" Family Structure.

The Table: 4.16 below shows the main effects of Family Structures on education attainment considering all the background household variables of Family size, SES, and Educational Level for head for Female-headed households.

Table 4.16

Full Model of Family Structures considering Family Size, Head Educational Level and Social Economic Status of a household for Female headed households

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	-4.296	.2616	-4.808	-3.783	269.707	1	.000	.014	.008	.023
ever married]	.405	.2637	-.112	.922	2.358	1	.125	1.499	.894	2.514
rried monogamous]	.433	.2485	-.054	.920	3.042	1	.081	1.542	.948	2.510
rried polygamous]	.789	.2645	.270	1.307	8.890	1	.003	2.200	1.310	3.695
dowed]	1.060	.2515	.567	1.553	17.769	1	.000	2.887	1.763	4.727
orced]	.894	.3099	.287	1.502	8.322	1	.004	2.445	1.332	4.489
parated]	0[a]									
ate]	-2.578	.1545	-2.881	-2.275	278.553	1	.000	.076	.056	.103
mary]	-2.634	.1398	-2.908	-2.360	354.799	1	.000	.072	.055	.094
condary]	-1.924	.1326	-2.184	-1.665	210.624	1	.000	.146	.113	.189
llege]	-.988	.1319	-1.247	-.730	56.118	1	.000	.372	.287	.482
iversity]	0[a]									
or]	-2.292	.2017	-2.687	-1.896	129.047	1	.000	.101	.068	.150
dle class]	-.986	.0934	-1.170	-.803	111.555	1	.000	.373	.311	.448
ch]	0[a]	.	.	.	.	.	.	1	.	.
sehold Size	-.002	.0030	-.008	.004	.361	1	.548	.998	.992	1.004
le)	1[b]									

The "Never Married" Family Structure having the least difference increase of .405 on the log count of undergraduates echos my earlier presumption that the reason why "Never Married" have had comparatively higher increase on mean log count of undergraduates on all the other occasions compared to "Separated" Family Structure is because of economically stable women who have chosen to leave independent and never get married. "Widowed" Family Structure has the most increase in this Model of 1.060 on mean log count of undergraduates compared to "Separated" Family Structure showing the resilience mentioned earlier about widowed women even without due advantage of income or education. Family size is not significant in predicting counts of undergraduates when all the variables are included.

The Table 4.17 below show main effect Model results for Male headed households.

Table: 4.17

Main effects model for Family Structures for Male Headed households

ameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
ercapt)	-8.187	.5774	-9.319	-7.056	201.100	1	.000	.000	8.971E-5	.001
ver Married]	2.302	.5837	1.158	3.446	15.553	1	.000	9.995	3.183	31.380
rried Monogamous]	1.927	.5776	.795	3.059	11.133	1	.001	6.871	2.215	21.315
rried Polygamous]	1.412	.5823	.271	2.553	5.879	1	.015	4.104	1.311	12.848
dowed]	1.968	.5951	.802	3.134	10.936	1	.001	7.157	2.229	22.976
orced]	-23.622	1.1459 E5	-224609.665	224562.421	.000	1	1.000	5.510E-11	.000	.[a]
parated]	0[b]	.	.	.	.	.	.	1	.	.
ale)	1[c]									

The Table: 4.18 below shows the main effects of Family Structures on education attainment considering all the background household variables of Family size, SES, and Educational Level for head for Male headed households.

Table: 4.18

ull Model of Family Structures considering Family Size, Head Educational Level and Social Economic Status f a household for Male headed households

meter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
rcept)	-4.188	.5780	-5.321	-3.056	52.514	1	.000	.015	.005	.047
r married]	1.579	.5839	.435	2.724	7.314	1	.007	4.851	1.545	15.236
ried monogamous]	1.375	.5777	.243	2.507	5.664	1	.017	3.955	1.275	12.272
ried polygamous]	1.347	.5825	.205	2.488	5.345	1	.021	3.845	1.228	12.042
owed]	1.834	.5952	.667	3.000	9.491	1	.002	6.257	1.949	20.089
rced]	.000	.	.	.	.	.	.	1.000	.000	.000
arated]	0^b	.	.	.	.	.	.	1	.	.
rate]	-3.578	.0993	-3.773	-3.383	1297.836	1	.000	.028	.023	.034
nary]	-4.009	.0635	-4.134	-3.885	3980.381	1	.000	.018	.016	.021
ondary]	-3.051	.0510	-3.151	-2.951	3580.038	1	.000	.047	.043	.052
ge]	-1.725	.0451	-1.814	-1.637	1465.569	1	.000	.178	.163	.195
ersity]	0^b									
]	-2.030	.1763	-2.376	-1.685	132.535	1	.000	.131	.093	.186
class]	-.978	.0407	-1.058	-.898	577.690	1	.000	.376	.347	.407
	0^b									
hold Size	-.004	.0028	-.010	.001	2.575	1	.109	.996	.990	1.001
)	1^c									

"Married polygamous" have the least increase on mean log count of undergraduates for both the full and Main effect Models as other register a leap in the expected mean log count of undergraduates compared to "Separated" Family Structure. This implies that all Male headed

Family Structures will have better prospects of children attaining education compared to "Separated" Family Structure.

The Table: 4.19 below show Zero-inflated Poisson Regression Model results for the Main effect of Family Structures on undergraduate counts.

Table: 4.19

ZIP Main Effects Model for Family Structures						
Poisson part				Binomial part		
	Parameter Estimate	Std. Error	Pr(>|z|)	Parameter Estimate	Std. Error	Pr(>|z|)
Intercept	-1.3320	0.1852	6.45e-13	4.0685	0.1794	< 2e-16
Married monogamous	-0.5362	0.1950	0.00597	-1.1907	0.1899	.63e-10
Married polygamous	-0.4376	0.2803	0.11850	-0.6137	0.2757	0.0260
Widowed	-0.5603	0.2738	0.04070	-0.6652	0.2699	0.0137
Divorced	0.6890	0.4344	0.11268	0.9706	0.4131	0.0188
Separated	-4.8091	1.8678	0.01003	-5.1064	7.0591	0.4694
Log-likelihood					-2.586e+04	
DF					12	

The Table: 4.20 below show Zero-inflated Poisson Regression Model results for the full Model for the effect of Family Structures on undergraduate counts.

Table: 4.20

ZIP Main Effects Model for Family Structures controlling for family size, Social economic status, and Educational level of head						
Poisson part				Binomial part		
	Parameter Estimate	Std. Error	Pr(>|z|)	Parameter Estimate	Std. Error	Pr(>|z|)
Intercept	-4.658343	0.761610	9.57e-10	3.736217	0.793172	2.47e-06
Married monogamous	0.924666	0.117236	3.09e-15	0.273358	0.169610	0.107030
Married polygamous	0.313258	0.188376	0.09632	-0.749043	0.249456	0.002676
Widowed	0.563957	0.182684	0.00202	-0.774292	0.234780	0.000974
Divorced	0.931014	0.320135	0.00364	0.579046	0.425465	0.173522
Separated	-0.759360	0.542102	0.16128	-0.624939	0.674549	0.354210
Family size	0.104694	0.002791	< 2e-16	-0.009228	0.001711	6.97e-08
Middle class	1.188942	0.741386	0.10879	0.005001	0.762768	0.994769
Rich	1.968907	0.743142	0.00806	-0.079966	0.766510	0.916911
Primary	-0.192714	0.266058	0.46886	-0.008856	0.269975	0.973830
Secondary	-0.128761	0.256798	0.61608	-0.844375	0.263306	0.001342
College	-0.301162	0.256133	0.23967	-2.503289	0.276019	< 2e-16
University	-0.162255	0.235937	0.49164	-16.148629	47.263428	0.732597
Log-likelihood					-1.991e+04	
DF					28	

The results for both Models agree with previous Poisson regression results. All the coefficients are now negative for the Main effect Model with "Never married" Family Structure as the reference category. This implies that effect of all the Family Structures leads to a decrease in the mean log count of undergraduates when compared to "Never married" Family Structure. "Separated" Family Structure has the least coefficient of -4.8091. This agrees with my Poisson results which showed "Never married" Family Structure as the one with the highest expected increase on mean log count of undergraduates while comparing with "Separated" Family Structure.

5 Chapter 5: CONCLUSION AND DISCUSSION

5.1 Discussion

By regressing counts of undergraduates from different Family structures captured during the 2009 Kenya Housing and population census, this study explored if indeed significant differences exists between the numbers coming from the various Family structures. Pair wise comparison of estimated means of undergraduates from each Family Structure showed significant differences between Family Structures which is consistent with previous studies though the other studies had mainly focused on intact versus non-intact Family Structures.

Previous studies have favoured intact Family Structure; that is "Married monogamous" but surprise result for this study is that "Never Married" has had the largest coefficient in estimating number of undergraduates. This is an indication that as we see more effort being channeled towards woman empowerment in Kenya, more and more women are choosing to remain single leading to a steep raise in families whose heads have never married. These are mainly comprised of women who have good education and career and are in a position to comfortably take care of their families to an extent that they are able to provide for what was arguably a preserve of children from intact families (Married monogamous). They are able to dedicate time for their children or employ qualified staff who takes care of their children in their absence. As such, conditions that have all along being thought to favour a child from "Married monogamous" Family Structure in excelling in education are being replicated in this "Never married" Family Structures and bringing in a new shape in how children are currently performing in education circles.

The other reason why this turnaround is being witnessed could be because of the dwindling interaction that prevails currently between parents from "Married monogamous" and their children as they try to cope with hard economic situation that has prevailed in Kenya for now close to two decades. Parents are forced to leave their homes early and return late for almost all days of the week. As such the comparative advantages children from "Married monogamous" Family Structures were presumed to have as compared to children from other Family Structures no longer measures up to what it used to be.

Most affected individuals are individuals from "Separated" Family Structures. "Married polygamous", "Divorced" and "Widowed" Family structures, have also not been spared but not as severely as "Separated" Family Structures. The reason is because unlike in these other 3 Family Structures, separation in Kenyan families doesn't come along with some form of support. Traditionally, widows always receive some form of support from extended family members not forgetting the taking over of any property that was left by the spouse. On the other hand divorce mostly occurs in those families that are relatively well off and in most cases it ends up with each spouse having some form of entitlement over whatever family resource that is there. There is also some support for "Married polygamous" families by virtue of both spouse being there. All this explains why children from "Separated" Family Structures are most affected considering that their single parents struggle with no support from any quarter and also probably both the parents and children are likely to be traumatized by the separation.

It is evident that children from Male-Headed or Female-headed households are affected differently after all the Family Structures registered negative coefficients for Female-headed households except the "Never married" Family Structure. As such children from Family Structures that are Female-headed can be said to be affected more than counterparts from Male-headed Family structures except for those from "Never married" family structure which has been discussed. By virtue of the fact that men heading households are in a better position to cater for their families than women, this explains why their children are not affected as much as those from Female-headed households.

This study would have better explored the relationship between Family Structures and education attainment of children with the availability of more information like religion and quality of schools which was not available from data provided by Kenya National Bureau of Statistics.

5.2 Conclusion

From my discussion it can be concluded that children's education attainment is highly influenced by the Family structure they are coming from. Most affected are children from separated Family structures and if there is to be any intervention more focus should be on children from this Family Structure.

Affected on equal measure are children from Married polygamous, Divorced and Widowed Family Structures. This means that children from these Family Structures should be viewed as suffering the same fate and same approach of mitigation should be thought of when it comes to mitigating efforts.

It has also be seen that probably because of prevailing economic situation in Kenya, children from intact Family Structure (Married monogamous) are no longer advantaged on educational front as previous studies have always purported but children from single families of women who have never married are having it better with their economically stable parent. Thus focus on children from economically stable never married single mothers should not be as much.

The fate of children who hail from Female-headed households should also be urgently looked into except for those from never married economically stable women. From my findings number of undergraduates from female-headed households is seen to be on the decline for all Family Structures except the never married Family Structure.

5.3 Recommendation

Having found that significant differences exists between number of children who complete undergraduate studies coming from different Family Structures in Kenya, it is imperative for the Ministry of Education to reconsider designing programmes that take due regard of the kind of family structure a child is coming from. Over the years it has been presumed that only children with disability need special attention in devising ways of imparting them with education. This consideration of special need should not just stop at disability only but should instead extend further to incorporate family background considerations of the children who are coming to schools.

Special offices should be set up in every school aimed at keeping a close eye on children who are known to come from non-intact families structures and monitoring children's performance with a view of following up on any unusual downward trends in performances orchestrated by a child's prevailing family background. These offices should have qualified counselors who will be seeking audience with concerned children or their existing parents at appropriate moments. Going by the findings of this study, keen eye should always be kept on children from "Separated" Family Structures even as the others are equally being monitored.

A timely word of encouragement and guidance can go a long way in turning a children's future around especially on matters education.

6 Chapter 6: REFERENCES

˙Ilknur ¨Ozmen1, F. F. (2006). Count Regression Models with an Application to Zoological Data Containing Structural Zeros. *Journal of Data Science 5(2007), 491-502)* .

'02, M. D. (2002). The Effects of Family, Social and Background Factors on Children's Educational Attainment. *Illinois Wesleyan University Economics Department Honors Projects Paper 8.*

A. Colin Cameron, P. K. (1999). *Essentials of Count Data Regression.*

Abby Córdova, P. (2008). Methodological Note:Measuring Relative Wealth using Household Asset Indicators. *AmericasBarometer Insights: 2008 (No.6)** .

Abilla, C. M. (1998). *A case study of special education in Kenya.* Nairobi: UNESCO.

Agresti, A. (1982). *Categorical Data Analysis.* London: Wiley Series.

Buchmann, C. (2000). Family Structure,Parental Perceptions and Child labour in Kenya: What factors determine who is enrolled in school? *Social forces, Vol.78, No.4* , 1349-1378.

C. E. Rose, S. W. (2012). On the Use of Zero-Inflated and Hurdle Models for Modeling Vaccine Adverse Event Count Data . *Journal of Biopharmaceutical Statistics* .

Cameron, A. C. (2000). *Regression Analysis of Count data.* London: Wiley Series.

Carter, W. Y. (1999). The Effects of Changing Family Structures on Higher Education for Black and WhiteAmerican Cohorts: 1908-1969. *CDE Working Paper No. 96-22* .

Hugo NAYA, J. I.-M.-M. (2008). A comparison between Poisson and zero-inflated Poisson regression models with an application to number of black spots in Corriedale sheep. *Genet. Sel. Evol. 40 (2008) 379–394* .

Lloyd, C. B. (2001). Determinants of Educational Attainment Among Adolescents in Egypt: Does School Quality Make a Difference? *Policy Research Division Working Paper no. 150.New York: Population Council.*

LUIGI ALDIERI, C. P. (2005). *AN INVESTIGATION OF THE RELATION BETWEEN THE NUMBER OF CHILDREN AND EDUCATION IN ITALY.* Napoli: JEL Classification: I21, J13, J24.

Mustafa, H. E. (2008). Socioeconomic Determinants of Infant Mortality in Kenya: Analysis of Kenya-DHS 2003. *Humanities and Social Sciences Volume 2, Issue 2, 2008* .

Professor Charles Desforges, A. A. (2003). The Impact of Parental Involvement, Parental Support and Family Education on Pupil Achievements and Adjustment. *Research Report RR433* .

Quincy Thomas Stewart, S. S. (2005). *Are Assets a Valid Proxy for Income? An Analysis of Socioeconomic Status and Child Mortality in South Africa.* Bloomington: Indiana University.

Regression Analysis of Count Data. (2001). *Jour. Ind. Soc. Ag. Statistics* , 220·231.

Richard Berk, J. M. (2007). *Overdispersion and Poisson Regression.* Department of Statistics, University of Pennsylvania.

Ridout, M. (December 1998). Models for count data with many zeros. *International Biometric Conference.* Cape Town.

SEEMA VYAS, L. K. (2006, October 9). Constructing socio-economic status indices:how to use principal components analysis. *London School of Hygiene and Tropical Medicine, London, UK* .

Shin, J. (2011). *Mixed-effect and mixed distribution models for count data with application educational research data.*

Tara G. Martin, B. (2005). Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. *Ecology Letters, (2005) 8: 1235–1246* .

Teresa Abada, F. H. (2008). Group Differences in Educational Attainment Among the Childrenof Immigrants. *11F0019M No. 308* .

Thomas N O Achia, A. W. (2010). A Logistic Regression Model to Identify Key Determinants of Poverty Using Demographic and Health Survey Data. *European Journal of Social Sciences – Volume 13, Number 1* .

Uwaifo, V. O. (2008). The Effects of Family Structure and Parenthood on the Academic Performance of Nigerian University Students. *Stud Home Comm Sci, 2(2): 121-124* .

Wambugu, A. (2002). Family Background, Education and Earnings in Kenya. *Working Paper s in Economics no 76* .

7 Chapter 7: APPENDEXIS

7.1 Codes for counting number of undergraduates and usual members of households and combining some variables

```
select* from (select distinct
bar.barcode,
isnull(grad.undergraduates, 0) as
undergraduates,
   isnull(size.householdsize, 0) as
householdsize
 from dat bar
 left outer join

  (select barcode, count(*) as
undergraduates
     from dat
     where p41='18'
      and P14='1'
      and P10 <> '1'
     group by barcode) grad on
grad.barcode = bar.barcode
 left outer join

   (select barcode, isnull(COUNT(*), 0)
as householdsize
     from dat
     where P14 = '1'
     group by BARCODE) size on
size.BARCODE = bar.BARCODE)c

 left join
  (select [BARCODE]
    ,[HHNO]
    ,[LINE_NUMBER]
    ,[P10]
    ,[P11]
    ,[P12]
    ,[P14]
    ,[P17]
    ,[P41]
    ,[P42]
    ,[H17_1]
    ,[H17_2]
    ,[H17_3]
    ,[H17_4]
    ,[H17_5]
    ,[H17_6]
    ,[H17_7]
    ,[H17_8]
    ,[H17_9]
    ,[H17_10]
    ,[H17_11]
    ,[H20]
    ,[H21]
    ,[H22]
    ,[H23]
    ,[H24]
    ,[H25]
    ,[H26]
    ,[H27]
    ,[H28_1]
    ,[H28_2]
    ,[H28_3]
```

```
compute New_P41=P41.
recode New_P41 (97 21 25 96 22
26 0=0) (23 1 2 3 4 5 6 7 8 9=1)
(15 24 10 11 12=2) (17 13 14
16=3) (18 19 20=4) (98 99=99).
val lab New_P41 0 "Illiterate" 1
"Primary" 2 "Secondary" 3
"College" 4 "University".
exe.
var lab
New_P41  Highest level of
education.
compute Illiterate =
(New_P41=0).
val lab Illiterate  0 "No" 1 "Yes".
exe.
var lab
Illiterate   Whether an Illiterate
individual in the HH.

compute Primary = (New_P41=1).
val lab Primary 0 "No" 1 "Yes".
exe.
var lab
Primary Whether a person whose
level of education is primary is in
the HH.

compute Secondary =
(New_P41=2).
val lab Secondary 0 "No" 1 "Yes".
exe.
var lab
Secondary   Whether a person
whose level of education is
secondary is in the HH.

compute College = (New_P41=3).
val lab College  0 "No" 1 "Yes".
exe.
var lab
College   Whether a person whose
level of education is college is in
the HH.

compute University =
(New_P41=4).
val lab University 0 "No" 1 "Yes".
exe.
var lab
University  Whether a person
whose level of education is
university is in the HH.
compute CorrugatedIronsheets  =
(H21=1).
val lab CorrugatedIronsheets  0
"No" 1 "Yes".
exe.
var lab
```

```
compute Purchased =
(H20=1).
val lab Purchased  0 "No" 1
"Yes".
exe.
var lab
Purchased  Purchased
house.

compute Constructed =
(H20=2).
val lab Constructed 0 "No" 1
"Yes".
exe.
var lab
Constructed  Constructed
House.

compute Inherited =
(H20=3).
val lab Inherited 0 "No" 1
"Yes".
exe.
var lab
Inherited  Inherited House.


compute Government =
(H20=4).
val lab Government 0 "No"
1 "Yes".
exe.
var lab
Government  Government
House.

compute LocalAuthority =
(H20=5).
val lab LocalAuthority 0
"No" 1 "Yes".
exe.
var lab
LocalAuthority  Local
authority House.

compute Parastatal =
(H20=6).
val lab Parastatal 0 "No" 1
"Yes".
exe.
var lab
Parastatal  Parastatal
House.

compute PrivateCompany =
(H20=7).
val lab PrivateCompany 0
"No" 1 "Yes".
exe.
```

```
val lab PipedDwelling 0 "No" 1
"Yes".
exe.
var lab
PipedDwelling   Water piped
into dwelling unit.

compute Piped = (H24=11).
val lab  Piped 0 "No" 1 "Yes".
exe.
var lab
 Piped   Water piped into
compound.

compute Jabia = (H24=12).
val lab Jabia  0 "No" 1 "Yes".
exe.
var lab
Jabia    Water from Jabia .

compute Rain = (H24=13).
val lab Rain 0 "No" 1 "Yes".
exe.
var lab
Rain Rain Water harvesting.

compute WaterVendor =
(H24=14).
val lab WaterVendor 0 "No" 1
"Yes".
exe.
var lab
WaterVendor  Water from
Vendors.

compute OtherSources =
(H24=15).
val lab OtherSources 0 "No" 1
"Yes".
exe.
var lab
OtherSources   From Other
Sources of Water.

compute MainSewer =
(H25=1).
val lab  MainSewer 0 "No" 1
"Yes".
exe.
var lab
 MainSewer   Houses with
Main Sewer disposal.
compute Lake = (H24=3).
val lab Lake 0 "No" 1 "Yes".
exe.
var lab
Lake Water from Lake.
```

```
    ,[H28_4]
    ,[H28_5]
    ,[H28_6]
    ,[H28_7]
    ,[H28_8]
    ,[H28_9]
    ,[H28_10]
    ,[H28_11]
    ,[H28_12]
    ,[H28_13]
    ,[H28_14]
     from dat
  where P10='1')d on
d.BARCODE=c.BARCODE

compute Othertypes = (H21=9).
val lab Othertypes 0 "No" 1 "Yes".
exe.
var lab
Othertypes  Other types of roofs.

compute Stone = (H22=1).
val lab Stone 0 "No" 1 "Yes".
exe.
var lab
Stone  Stone walled Houses.

compute Brick = (H22=2).
val lab Brick 0 "No" 1 "Yes".
exe.
var lab
Brick Brick or block walled Houses.


compute MudWood = (H22=3).
val lab MudWood 0 "No" 1 "Yes".
exe.
var lab
MudWood   Houses with Mud or Wood
walls.

compute MudCement  = (H22=4).
val lab MudCement 0 "No" 1 "Yes".
exe.
var lab
MudCement   Houses with Mud
orCement walls.


compute Otherwalls = (H22=9).
val lab Otherwalls 0 "No" 1 "Yes".
exe.
var lab
Otherwalls  Other types of walls.

compute Cemented = (H23=1).
val lab Cemented 0 "No" 1 "Yes".
exe.
var lab
Cemented   Houses with Cemented
floor.

compute Tilled = (H23=2).
val lab  Tilled 0 "No" 1 "Yes".
exe.
var lab
CorrugatedIronsheets   Iron Sheet
Houses.

compute Tiled = (H21=2).
val lab Tiled 0 "No" 1 "Yes".
exe.
var lab
Tiled  Tiled roof Houses.


compute Concrete = (H21=3).
val lab Concrete 0 "No" 1 "Yes".
exe.
var lab
Concrete   Houses with Concrete
slabs.

compute Asbestos  = (H21=4).
val lab Asbestos 0 "No" 1 "Yes".
exe.
var lab
Asbestos    Houses with Asbestos
Sheets roofs.

compute Grass = (H21=5).
val lab Grass 0 "No" 1 "Yes".
exe.
var lab
Grass   Grass thatched Houses.

compute Makuti = (H21=6).
val lab Makuti 0 "No" 1 "Yes".
exe.
var lab
Makuti  Makuti thatched Houses.

compute Tin = (H21=7).
val lab Tin  0 "No" 1 "Yes".
exe.
var lab
Tin    Houses with Tin roofs.

compute Mud = (H21=8).
val lab Mud 0 "No" 1 "Yes".
exe.
var lab
Mud House with Mud or Dung
roofs.
compute Wood = (H22=5).
val lab Wood 0 "No" 1 "Yes".
exe.
var lab
Wood   Wood walled Houses.

compute Ironsheets = (H22=6).
val lab  Ironsheets 0 "No" 1 "Yes".
exe.
var lab
 Ironsheets   Ironsheets walled
Houses.

compute Grasswalled = (H22=7).
val lab Grasswalled  0 "No" 1
"Yes".
exe.
var lab
var lab
PrivateCompany  Private
company's House.

compute Individual =
(H20=8).
val lab Individual  0 "No" 1
"Yes".
exe.
var lab
Individual   individual's
House.

compute FaithbasedNGO =
(H20=9).
val lab FaithbasedNGO 0
"No" 1 "Yes".
exe.
var lab
FaithbasedNGO  NGO's
House.

compute Otherform =
(H20=10).
val lab Otherform 0 "No" 1
"Yes".
exe.
var lab
Otherform  Other type of
tenure.
compute Earth = (H23=4).
val lab Earth 0 "No" 1 "Yes".
exe.
var lab
Earth Houses with Earth
floor.

compute Otherfloors =
(H23=5).
val lab Otherfloors 0 "No" 1
"Yes".
exe.
var lab
Otherfloors  Other types of
floors.

compute Pond = (H24=1).
val lab Pond 0 "No" 1 "Yes".
exe.
var lab
Pond  Water from Pond.

compute Dam = (H24=2).
val lab Dam 0 "No" 1 "Yes".
exe.
var lab
Dam  Water from Dam.

compute Lake = (H24=3).
val lab Lake 0 "No" 1 "Yes".
exe.
var lab
Lake Water from Lake.


compute Stream = (H24=4).
compute Stream = (H24=4).
val lab Stream 0 "No" 1 "Yes".
exe.
var lab
Stream  Water from Stream.

compute Borehole  = (H24=9).
val lab Borehole 0 "No" 1
"Yes".
exe.
var lab
Borehole    Water from
Borehole.

compute PipedDwelling =
(H24=10).
compute SepticTank =
(H25=2).
val lab SepticTank  0 "No" 1
"Yes".
exe.
var lab
SepticTank    Houses with
Septic Tank disposal.

compute CessPool = (H25=3).
val lab CessPool 0 "No" 1
"Yes".
exe.
var lab
CessPool Houses with Cess
Pool disposal.
compute VIPLatrine = (H25=4).
val lab VIPLatrine 0 "No" 1
"Yes".
exe.
var lab
VIPLatrine  Houses with VIP Pit
Latrine waste disposal.

compute PitLatrine = (H25=5).
val lab  PitLatrine 0 "No" 1
"Yes".
exe.
var lab
 PitLatrine   Houses with Pit
Latrine covered waste
disposal.

compute PitLatrineUncovered
= (H25=6).
val lab PitLatrineUncovered  0
"No" 1 "Yes".
exe.
var lab
PitLatrineUncovered    Houses
with Pit Latrine Uncovered
waste disposal              .

compute Bucket = (H25=7).
val lab Bucket 0 "No" 1 "Yes".
exe.
var lab
Bucket Houses with Bucket
Latrine waste disposal.
```

```
Tilled  Houses with Tilled floor.

compute Wooded = (H23=3).
val lab Wooded  0 "No" 1 "Yes".
exe.
var lab
Wooded    Houses with wood floor .

compute Earth = (H23=4).
val lab Earth 0 "No" 1 "Yes".
exe.
var lab
Earth Houses with Earth floor.

compute Otherfloors = (H23=5).
val lab Otherfloors 0 "No" 1 "Yes".
exe.
var lab
Otherfloors  Other types of floors.

compute Pond = (H24=1).
val lab Pond 0 "No" 1 "Yes".
exe.
var lab
Pond  Water from Pond.

compute Dam = (H24=2).
val lab Dam 0 "No" 1 "Yes".
exe.
var lab
Dam  Water from Dam.

compute Solar = (H26=7).
val lab Solar 0 "No" 1 "Yes".
exe.
var lab
Solar  Houses with Solar as main source of cooking fuel.

compute OtherFuel = (H26=8).
val lab OtherFuel 0 "No" 1 "Yes".
exe.
var lab
OtherFuel  Houses with Other sources as main source of cooking fuel.

Grasswalled   Grass walled Houses .

compute Tinned = (H22=8).
val lab Tinned 0 "No" 1 "Yes".
exe.
var lab
Tinned House with Tin walls.
compute OtherMethods = (H25=9).
val lab OtherMethods 0 "No" 1 "Yes".
exe.
var lab
OtherMethods   Houses with other methods of waste disposal.

compute Electricity = (H26=1).
val lab Electricity  0 "No" 1 "Yes".
exe.
var lab
Electricity    Houses with Electricity as main source of cooking fuel.

compute Paraffin = (H26=2).
val lab Paraffin 0 "No" 1 "Yes".
exe.
var lab
Paraffin Houses with Paraffin as main source of cooking fuel.
compute FirewoodLighting = (H27=6).
val lab FirewoodLighting 0 "No" 1 "Yes".
exe.
var lab
FirewoodLighting Houses with Firewood Lighting as main source of cooking lighting.
compute Charcoal = (H26=6).
val lab Charcoal 0 "No" 1 "Yes".
exe.
var lab
Charcoal Houses with Charcoal as main source of cooking fuel
compute Biogas = (H26=4).
val lab  Biogas 0 "No" 1 "Yes".
exe.
var lab
 Biogas   Houses with Biogas as main source of cooking fuel.

val lab Stream 0 "No" 1 "Yes".
exe.
var lab
Stream  Water from Stream.
compute SolarLighting = (H27=7).
val lab SolarLighting 0 "No" 1 "Yes".
exe.
var lab
SolarLighting  Houses with Solar Lighting as main source of cooking lighting.

compute OtherLighting = (H27=8).
val lab OtherLighting 0 "No" 1 "Yes".
exe.
var lab
OtherLighting   Houses with Other sources as main source of lighting.

compute Exotic_cattle = (H17_1 >= 1).
val lab Exotic_cattle 0 "No" 1 "Yes".
exe.
var lab
Exotic_cattle Whether exotic cattle was in the HH.
compute TinLamp = (H27=4).
val lab  TinLamp 0 "No" 1 "Yes".
exe.
var lab
 TinLamp   Houses with Tin Lamp as main source of cooking lighting.

compute GasLamp = (H27=5).
val lab GasLamp 0 "No" 1 "Yes".
exe.
var lab
GasLamp    Houses with Gas Lamp as main source of cooking lighting

compute Bush = (H25=8).
val lab Bush 0 "No" 1 "Yes".
exe.
var lab
Bush  Houses with Bush waste disposal.
compute ElectricityLight = (H27=1).
val lab ElectricityLight  0 "No" 1 "Yes".
exe.
var lab
ElectricityLight    Houses with Electricity as main source of lighting.

compute PressureLamp = (H27=2).
val lab PressureLamp 0 "No" 1 "Yes".
exe.
var lab
PressureLamp Houses with Pressure Lamp as main source of lighting.

compute Lantern = (H27=3).
val lab Lantern 0 "No" 1 "Yes".
exe.
var lab
Lantern  Houses with Lantern as main source of cooking lightingl.
compute Firewood = (H26=5).
val lab Firewood  0 "No" 1 "Yes".
exe.
var lab
Firewood    Houses with Firewood as main source of cooking fuel.
compute LPG = (H26=3).
val lab LPG 0 "No" 1 "Yes".
exe.
var lab
LPG  Houses with LPG as main source of cooking fuel.
```

7.2 Codes for breaking down categorical variables for construction of Social Economic Status of households

```
Compute Tiles=$sysmis.
IF (Tiled=0 & Asbestos=0) Tiles=0.
IF (Tiled=0 & Asbestos=1) Tiles=1.
IF (Tiled=1 & Asbestos=0) Tiles=1.
IF (Tiled=1 & Asbestos=1) Tiles=1.
EXECUTE .
Compute Latrines=$sysmis.
IF (PitLatrine=0 & PitLatrineUncovered=0) Latrines=0.
IF (PitLatrine=1 & PitLatrineUncovered=0) Latrines=1.
IF (PitLatrine=0 & PitLatrineUncovered=1) Latrines=1.
IF (PitLatrine=1 & PitLatrineUncovered=1) Latrines=1.
EXECUTE .
fre Latrines.
Compute GreenLighting=$sysmis.
IF (PressureLamp=0 & GasLamp=0 & SolarLighting=0)
GreenLighting=0.
IF (PressureLamp=1 & GasLamp=0 & SolarLighting=0)
GreenLighting=1.
IF (PressureLamp=0 & GasLamp=1 & SolarLighting=0)
GreenLighting=1.
IF (PressureLamp=0 & GasLamp=0 & SolarLighting=1)
GreenLighting=1.
IF (PressureLamp=1 & GasLamp=1 & SolarLighting=1)
GreenLighting=1.
IF (PressureLamp=0 & GasLamp=1 & SolarLighting=1)
GreenLighting=1.
IF (PressureLamp=1 & GasLamp=0 & SolarLighting=1)
GreenLighting=1.
IF (PressureLamp=1 & GasLamp=1 & SolarLighting=0)
GreenLighting=1.
EXECUTE .
fre GreenLighting.
Compute ExpLamps=$sysmis.
IF (PressureLamp=0 & GasLamp=0 ) ExpLamps=0.
IF (PressureLamp=1&  GasLamp=0 ) ExpLamps=1.
IF (PressureLamp=0 & GasLamp=1 ) ExpLamps=1.
IF (PressureLamp=1 & GasLamp=1 ) ExpLamps=1.
EXECUTE .

Compute PipedWater=$sysmis.
IF (Piped=0 & PipedDwelling=0 ) PipedWater=0.
IF (Piped=1 & PipedDwelling=0) PipedWater=1.
IF (Piped=0 & PipedDwelling=1 ) PipedWater=1.
IF (Piped=1 & PipedDwelling=1 ) PipedWater=1.
EXECUTE .
```

```
Compute IronRoofed=$sysmis.
IF (Corrugated=0 & Tin=0) IronRoofed=0.
IF (Corrugated=1 & Tin=0) IronRoofed=1.
IF (Corrugated=0 & Tin=1) IronRoofed=1.
IF (Corrugated=1 & Tin=1) IronRoofed=1.
EXECUTE .
Compute CommonFuel=$sysmis.
IF (Paraffin=0 & Charcoal=0) CommonFuel=0.
IF (Paraffin=1 & Charcoal=0) CommonFuel=1.
IF (Paraffin=0 & Charcoal=1) CommonFuel=1.
IF (Paraffin=1 & Charcoal=1) CommonFuel=1.
EXECUTE .
fre CommonFuel.
Compute IronAsbestos=$sysmis.
IF (CorrugatedIronsheets=0 & Asbestos=0 )
IronAsbestos=0.
IF (CorrugatedIronsheets=1 & Asbestos=0 )
IronAsbestos=1.
IF (CorrugatedIronsheets=0 & Asbestos=1 )
IronAsbestos=1.
IF (CorrugatedIronsheets=1 & Asbestos=1 )
IronAsbestos=1.
EXECUTE .
Compute Latrines=$sysmis.
IF (PitLatrine=0 & PitLatrineUncovered=0 ) Latrines=0.
IF (PitLatrine=1 & PitLatrineUncovered=0 ) Latrines=1.
IF (PitLatrine=0 & PitLatrineUncovered=1 ) Latrines=1.
IF (PitLatrine=1 & PitLatrineUncovered=1 ) Latrines=1.
EXECUTE .
FRE SES
/FORMAT=NOTABLE
/NTILES= 3
/STATISTICS=MAX MIN STDDEV MEAN
/ORDER= ANALYSIS.

Compute STATUS=$sysmis.
IF (SES<-0.2559) STATUS=1 OR
IF (SES>=-0.2559 & SES<2.8341) STATUS=2 OR
IF (SES>=2.8341 & SES<= 3.6472) STATUS=3.
EXECUTE .
```