UNIVERSITY OF NAIROBI

COLLEGE OF BIOLOGICAL AND PHYSICAL SCIENCES

SCHOOL OF MATHEMATICS

RESEARCH PROJECT IN MSC SOCIAL STATISTICS

MODELING OF HOUSEHOLD DECISION MAKING USING A TWO STAGE CLUSTER SURVEY DATA

BY

ERIC K. WAMITI

REG NUMBER I56/76592/09

Submitted to the School of Mathematics in partial fulfillment of the requirements for the degree of Master of Science in Social Statistics

Declaration

Declaration by the candidate

I **Eric K. Wamiti** hereby declare that this work has not previously been accepted in substance for any degree and is not being concurrently submitted to candidature for any degree in any other university.

E. K. Wamiti _____ 29/07/2011

Signature Date

Declaration by the Supervisor

This project is being submitted in partial fulfillment of the requirements of the Degree of Master of Science in Social Statistics with approval by university supervisor

Prof. M. MANENE

University of Nairobi

School of Mathematics

P.O Box 30197

Nairobi _____ 1/8/2011

Signature Date

Dedication

I dedicate this work to my family, my classmates and the lecturers of school of mathematics department of statistics and operations research.

Acknowledgements

First and foremost I am indebted to the School of mathematics department of statistics and operations research for the knowledge they have imparted to me in this two year's course, and especially professor M Manene, whose guidance and thorough discussions and suggestions has seen me this far. I wish to salute my fellow colleagues for the support and encouragement they have accorded to me.

Table of Contents

CHAPTER I

INTRODUCTION

1.1 Background of the study

Family is a fundamental building block of a society. A family performs several roles and serves as a source of stability, continuity and development of a nation. Family is a significant decision-making unit, deciding on a great number of issues that form components of the everyday life of a household.

Family decision making has been a central issue for researchers and scientists in the field of sociology. To evaluate the relationship between factors that influence decision making, it is important to start by studying family decision making from the household point of view.

Family financial management can be as well is used as an analysis of the inequality among families. It is a notion of anticipating family's position in a society.

This study uses two stage cluster sampling to study work, attitudes and spending in households from central province. Cluster sampling is economically justified when reduced costs can be used to overcome losses in precision. This is most likely to occur in a situation where population is concentrated in "natural" clusters (provinces, districts, locations etc.). This study conducted personal interviews of households in three district selected in central province by cluster sampling and three hundred and sixty eight households selected from the three districts by stratified sampling. It makes sense to randomly select a sample of districts (stage 1 of cluster sampling) and then interview house hold selected from the three districts by simple random sampling (stage 2 of cluster sampling). Using 2 stage cluster sampling, the interviewer could conduct many interviews in a single day. Dividing the population into groups and sampling from

a random subset of these groups (e.g. geographical locations) will decrease precision for a given sample size but often increase precision for a given cost.

In studying work we look at occupation and education of the respondent, in attitude we look at who makes key decisions in the house hold about various issues and how the respondents feel about gender issues like violence, freedom of financial decision making in the family, and in spending we look at ownership of various items in the household, the type of house the household owns, how much the household spend on items like food, social activities etc. Other demographic measures like age and place of birth are also included in the study.

The study uses 2 stage cluster sampling to estimate population parameters like the total, mean and proportion of various variables like gender, age education, levels of household control, levels of final decision makers in the household, expenditure on various items and ownership of various goods in the household. The study also looks at the ratio estimators for the population totals on household spending.

We also explore whether family decisions can be subjected to several variables among them income, education and various perception of the main decision maker. Several studies have demonstrated significant relationships between education, work, income and spending on key family decisions.

1.2 Statement of the problem.

In central province of Kenya, families fall under different structures. These structure depend on the various parameters like the education of the spouses, their levels of income, the family

decision maker, how family expenditure is controlled, wealth of the family, their social attitudes etc.

Family decision making indeed do much to explain the quality of livelihood depicted by the family. The question of interest is whether decisions are shared equally within the family.

Central province of Kenya have been highly challenged by issues of family mismanagement of available resources to an extent that there has been a felt increases in levels of poverty. To be able to explain this scenario we need to understand the parameters that determine different household financial management and decision making.

The overall research problem in this study is how households in central province make household decisions and what factors can be used to explain this.

1.3 Aims and objectives of the study.

The aim of this study was to determine through a two stage cluster sampling how families in central Kenya utilize their income, make family decisions and what factors influence the two processes.

1.4 The specific objectives of the study are to:

a) Examine how families in central Kenya spend their income and make important decisions.
b) Determine what factors influence decision making in the family.
c) Determine errors that result from analysing survey data as though it was a population data.

1.5 Significance of the study.

The family is a complex unit consisting of individuals with different thoughts, feelings, and ideas that can make communicating and decision-making challenging. The ability to make good decisions is a skill that, when developed over time, can help family members improve their skills. Effective group decision-making can reduce conflict between family and environment, as well as between family members.

The study aims at helping the sociologists to address the issue of the relationship between family decision making and financial management with other factors like education, age, income and attitudes. It is believed that equality between spouses in decision making is more likely to improve their quality of life and help them make decisions that would be of benefit to both the spouses and their children and have a more lasting and successful relationship.

The results of this analysis would help sociologists guide families to improve their quality of life through effective decision making.

CHAPTER II

LITERATURE REVIEW

Family economic status and decision making have been continuously discussed in the field of human ecology. These are some of the important key issues in the efforts of enhancing family and society well-being. Previous research have shown that demography variables affect family economic status and decision making, however, there is limited and unclear evidence on the relationship between decision making and family economic status.

In previous studies, little is known about differences in how married and cohabiting couples share their economic resources. Catharine Hobart Harner, (march, 2007) in 2001 panel of the Survey of Income and Program Participation (SIPP) compared the extent of household expense sharing among married and cohabiting couples focusing on gender specialization differences between married and cohabiting couples. The debates questioned the relative differences between married and cohabiting relationships. Multinomial logistic analysis suggested that relative resources indeed do much to explain who pays the majority of household expenses, but also found support for differences across family structure. The study found out that married couples are more likely than cohabiting couples to have a single male provider compared to other sharing arrangements. According to the study households with a child not biologically related to one partner are more likely to have a female provider, while households with biological children are more likely to have a male provider.

Married couples share economic resources through a variety of household allocative systems, including joint accounts, pooled income or "common pots," and separate money management practices.

In her study Catharine examined the extent to which both partners share expenses, both partners pay for expenses separately, the male partner pays household expenses, or the female partner pays household expenses in these different family arrangements. In her analysis, she took a careful consideration of relative resources of each partner across key characteristics, such as education, employment, and income, in order to more accurately assess relationship differences between married and cohabiting couples.

This study used data from wave three of the 2001 panel of the Survey of Income and Program Participation (SIPP). The 2001 panel of the SIPP began wave one with 35,100 interviews in eligible living quarters in 322 primary sampling units (PSUs). There were 89,141 people interviewed in wave one. An additional 14,100 individuals are estimated to have entered the sample during the two year period through births, marriages, and other reasons. Wave three and topical module three, the source of data on household expense sharing, contained approximately 71,280 individuals. The topical module in wave three of the survey included 27,401 households. The dependent variable was a household level variable constructed using information on whether each individual over age 15 in the household paid particular expenses with his/her own money. Descriptive statistics and nested multinomial logistic regression models were used to shed light on the relationships between marital status, children, relative resources, and couples' perceptions of responsibility for household expenses. Initial analyses focused on the characteristics of the populations across each of type of expense sharing: female provider, male provider, separate expenses, and sharing. In an effort to address major explanatory factors of couples' approaches to sharing household expenses, Catharine uses nested, multinomial logistic regression.

In the descriptive statistics majority of married and cohabiting couples were found to make joint contributions to household expenses (34 percent) or keep expenses separate (35 percent), while

just 3 percent were female payer/provider households, and 28 percent were male payer/provider households. This suggested that some form of mutual contribution to household expenses was most common. A larger proportion of male provider couples were married with biological children in household than any of the other three groups; cohabiting couples were more common among the other arrangements. The proportion of cohabiting couples among female payer/provider couples, couples making joint contributions, and couples keeping separate expense was largely similar.

Nine percent of female provider couples had a child not biologically related to one of the partners present, compared to between three and five percent in each of the other expense sharing arrangements. A higher proportion (37 percent) of female provider couples had experienced divorce than those in male provider or separate expense households, while a lower proportion of male provider households had experienced divorce than any other expense sharing type. Relative resource and socioeconomic differences were found across expense sharing arrangements. Female payer/provider couples had a larger proportion of women with a greater level of education than their partners, compared to male payer/providers and couples with separate-expense arrangements. Conversely, 34 percent of male payer/provider households had men with greater levels of education than their partners compared to 25 percent in joint contribution arrangements, 28 percent in separate expense arrangements, and 21 percent in female payer/provider households. In line with this, the female-to-male hours worked ratio in female payer/provider households were three times the mean of that across other modes of expense sharing. While women in female payer/provider relationships worked longer hours per week on average than women in other expense sharing arrangements, their male partners worked more hours on average than women in male payer/provider relationships. Finally, female and male

provider households had adjusted household income levels between 12 and 25 percent lower on average than separate and shared expense arrangements.

Married couples were fairly evenly distributed across male payer/provider, joint contribution and separate expense arrangements at 29, 33, and 35 percent respectively. The bulk of cohabiting households (44 percent) contributed jointly to expenses, followed by 35 percent in separate expense arrangements, 16 percent in male payer/provider households, and 4 percent in female payer/provider households. Twice as many married couples were two biological parent families compared to cohabiting couples, and, conversely, three times as many cohabiting as married couples had a child not biologically related to one partner in the household. There was no statistical difference between the percentages of married and cohabiting couples in which both partners reporting paying for all of their expense with their own money.

Female partners in cohabiting couples were found to have greater relative resources, on average, than female partners in married couples. A similar proportion of married and cohabiting couples had equal education. Women *and* men in cohabiting relationships worked more hours per week compared to married couples. Women in cohabiting relationships also worked more hours per week on average than their partners compared to married couples. A greater proportion of married female partners earned less than 25 percent of total household income, while a greater proportion of cohabiting female partners earned more than 50 percent of household income. A similar proportion of married and cohabiting female partners earned between 25 and 50 percent of household income, suggesting that both types of households included dual-earner arrangements.

Multinomial logistic regression used by Catharine indicated that married couples were more likely to have a male payer/provider than any other expense arrangement. Addition of controls

for the presence of children and previous divorce slightly reduced the coefficients associated with marital status. Presence of a child biologically related to both partners increased the odds that the household would be a male provider household compared to any of the other arrangements. The presence of a child not biologically related to a partner increased the odds that the household will be a female provider household rather than any of the other three arrangements. In addition, households with two-biological parents indicated the likelihood to be in any other arrangement compared to a male provider arrangement. Previous divorce for either partner increased the odds of a female payer/provider, couples contributing jointly, or couples reporting they both pay for expenses with their own money compared to male provider households. There were no statistical differences in the relationship with divorce across female provider, separate expense couples, and joint contribution expense arrangements.

(Mehdi Yadollahi, Laily Hj Paim, Mumtazah Othman, Turiman Suandi, 2009) described family economics as a determinant of the households and individuals level of living. Family economy has the potential for shifting the households and individuals level of living to a higher level via improving and meeting individuals and societal requirements.

Mehdi et al seeks to address the following question: (1) is there any relationship between management functions and family economic status? And (2) is there any difference of the effects of demographic variables on family economic status? The objectives of their study were: (1) to determine relationship between management functions and family economic status, and (2) to determine effect of demographic variables on family economic status. There were three measures of economic status that are used in this study: Income, expenditure or consumption and non-consumption, and wealth or asset.

The primary aim of this study was to identify factors affecting family economic status or to determine factors effecting family economic status. For this purpose, a multiple linear regression was used. The explanatory variables included in the model are age, level of education, and occupation and management functions in predicting family economic status.

Findings showed that the coefficient of adjusted multiple determinations is 0.264. Therefore, about 26.4 percent of the variation in family economic status is explained by all of explanatory variables (controlling, age of head household, level of education, occupation, and coordinating, organizing, directing and planning function). Level of education, age, occupation, and controlling were highly correlated. The impact of the correlation on the regression results was measured by deleting other variables (planning, coordinating, organizing, and directing) from the multiple linear regression analysis. Finally with 4 predictors, the coefficient of adjusted multiple a determination was 0.266: therefore, about 26.6 percent of the variation in family economic status was explained by all explanatory variables (controlling, age of head household, level of education, occupation).

Findings concerning the exact association between dependent variables and the individual independent variables require careful evaluation. In relation to the level of education and family economic status, results indicated that for every one unit increase in level of education there was an increase of 45.1 percent in the ability to predict the family economic status.

However, looking at the relationship between age and family economic status, the results indicated that for every one unit increase in age of the head of household there is an increase of 10.6 percent in the ability to predict family economic status. For every one unit increase in occupation there was a decrease of .276 units in the prediction of family economic status. For every one unit increase in controlling there was an increase of .106 units in predict of family

economic status. The model equation for the multiple linear regression was FES = -3.576+.451 level of education +.106 age -.276 occupation + .106 controlling where FES stands for Family Economic Status. The ANOVA test ($F\ (4, 385) = 36.19$, $P<0.000$) examines the degree to which the relationship between dependent variables and independent variables is linear. Since the F-test was significant, then relationship was found to be linear and therefore the model significantly predicted the family economic status.

Maria L. Sanz de Acedo Lizarraga and Baquedano (2007) studied the importance allocated to several factors in the decision process. Their main aim was to evaluate the importance that individuals allocate to the following aspects when they make decisions: uncertainty (six items), time/money constraints (eight items), information and goals (eight items), consequences of the decision (six items), motivation (five items), self-regulation (eight items), emotions (five items), cognition (six items), social pressure (seven items), and work pressure (five items). Factor analysis with varimax rotation and maximum likelihood extraction revealed the structure of 10 first-order factors that are integrated into 3 broader second-order factors: task (uncertainty, time/money pressure, information and goals, and consequences of decision), decision maker (motivation, self-regulation, cognition, and emotion), and environment (social pressure and work pressure). Each item was rated on a 9-point scale, with values ranging from 1 (*not at all important*) to 9 (*extremely important*).

Student's *t* test for independent samples revealed statistically significant differences between men and women in sex variable. The women allocated more importance than did the men to uncertainty, time/money constraints and consequences of the decision, the task factor, emotions, and social pressure. Conversely, men scored higher than women in information and goals,

motivation, and work pressure. No differences were found in cognition, self-regulation, and the environmental factor.

The analysis of variance of the age variable detected relevant differences in the three groups. The youths only achieved statistically higher scores than the other two groups in the variables emotion and social pressure. The adults revealed significant differences compared with the other groups in time/money constraints, information and goals, and work pressure, and the retired people only scored higher in uncertainty. Post hoc tests revealed statistically significant differences in the three age groups in uncertainty, time/money constraints, information and goals, emotion, and social pressure.

In many of the above reviews though data was sampled from population, no methods of error estimation or complex survey methods were used and hence data were treated as if it were population data.

The fact that survey data are obtained from units selected with complex sample designs needs to be taken into account in the survey analysis: weights need to be used in analyzing survey data and variances of survey estimates need to be computed in a manner that reflects the complex sample design.

Household surveys are commonly designed to produce estimates of population totals, population means, or simple ratios of totals or means. Examples of totals might be total population, men in the work force, women in the work force, or the number of children five years old or younger. Examples of means might be average income for persons in the work force, average income of women in the work force, and average income of men in the work force. Ratio estimates might be required to estimate the proportion of households with total income below the poverty level or the average household income for households whose principal wage earner is a female.

Descriptive reports may include standard errors of estimates or interval estimates based on those standard errors. Estimation of the standard errors requires an analysis that takes account of the household survey sample design. Interval estimates require not only the appropriate design-based estimates of standard errors. but also require knowledge of the degrees of freedom used in computation of the standard error estimates. These types of fairly simple descriptive statistics constitute the majority of the official statistics published to describe the results of household surveys.

In the analysis of complex survey data, comparisons usually involve t-tests. One basic issue involved in these tests is the calculation of degrees of freedom. A typical two-stage complex survey design has primary sampling units (PSUs) selected with probability proportional to size (PPS) independently within first-stage strata and secondary sampling units (SSUs) selected independently within PSUs using simple random sampling (SRS), (Allison M. Burns, Robert J. Morris, Jun Liu and Margaret Z. Byron, 2003).

When constructing confidence intervals for population parameters using survey data, it is usually assumed that survey estimates of mean and proportion are approximately normally distributed, so that the usual normal confidence intervals or, for small sample sizes n, t-distribution intervals with n-1 degrees of freedom are appropriate (Cochran 1977, p.27). For multistage sampling, the degree of freedom for t-intervals is typically equal to the number of sampled primary sampling units minus the correction, c, where c = 1 for unstratified cluster sampling and c equals the number of strata for stratified cluster sampling. (Korn and Graubard, 1999). The normality approximation may break down for estimate of proportions that are very close to 0 or 1. In these situations Korn and Graubard (1998; 1999, p.65) provide approximation to exact binomial confidence intervals that can be used with data from complex multistage sampling schemes.

Survey weights and statistical estimation based on those weights provide the link between the observations from a probability sample of households and summary measures or population parameters about the household population. The population of all households is sometimes called the target population or the universe. Without the application of both probability sampling and weighting, there is no supporting statistical theory to provide a link between the sample observations and the target population parameters.

CHAPTER III

METHODOLOGY

Secondary data on work attitude and spending (WAS survey) was used in this study. This data is available free from the official website for the work and attitude and spending (WAS) group of surveys. These are household surveys intended to be used by academics and campaigners. This is an ongoing project which began in 1992, aiming to understand why households behave as they do. It focuses on poor countries; the countries studied so far include India, Kenya, Brazil, Indonesia, Nigeria, Kenya, and Egypt.

John Simister commissioned these 'Work, Attitudes & Spending' (WAS) surveys, to give insights into household spending – what each household spends money on, and why. In particular, it focuses on whether the (usually male) head-of-household controls spending, or whether the (usually female) spouse has much control. Other aspects of power include what durable goods are owned, and who does the housework (husband, wife, or hired help). Possible determinants of male & female power are studied, including earnings; education; attitudes of respondent; and (in later WAS surveys) domestic violence. In each household studied, one person is interviewed (the respondent is also asked about other household members). Various demographic indicators are also included, such as birthplace.

In Kenya the survey was conducted in 2004 by SBO Research Ltd. (formerly called Strategic Business Opportunities Ltd.), SBO House, Jabavu Lane off Argwings Kodhek Rd, Hurlingham, Nairobi; contact address P.O. Box 10567, 00400 Nairobi, Kenya.

In conjunction with the data on WAS survey I requested data on projected number of households in Kenya in 2004 from Kenya National Bureau of Statistics. The data give the number of

provinces and the number of districts in each province. In each district the projected number of households is also given.

The WAS survey data consists of the entire households in the eight provinces of Kenya. Data from households sampled from Central province of Kenya was used in this study.

Central province has seven districts and only three districts were sampled. The sample selection was done as shown in the table below.

DISTRICT	NUMBER OF HOUSEHOLDS	HOUSEHOLDS SELECTED
KIAMBU	216,693	59
KIRINYAGA	143,966	
MURANGA	106,686	153
NYANDARUA	126,496	
NYERI	205,322	156
THIKA	196,412	
MARAGUA	111,332	
TOTAL	**1,106,909**	368

Table 1 Sample Selection

The information on the table indicates that two stage cluster sampling was done. The first sample was cluster sampling of three districts from seven districts in central province. The three districts constituted the primary sampling unit (PSU). From the three districts households were selected by simple random sampling. The 368 households, 59 from Kiambu, 153 from Muranga and 156 from Nyeri constituted the secondary sampling units (SSU).

The questionnaire on work, attitude and spending contained questions related to the three fields. These included among others, education, nature of employment, earnings, gender, age, spending, attitudes an opinions, place of birth, hours spent on paid employment, marital status, religion, ownership of certain goods like computers and television sets, type of house, how household expenses are done, who make important financial decisions and who controls household expenditure.

For two stage cluster sampling, clusters are selected at first by simple random sampling. Listed units in the selected clusters are then selected by simple random sampling at the second stage. The sampling fraction $\left[\frac{n_i}{N_i}\right]$may be the same or different at the second stage. In this case the sampling fraction is different.

A two stage cluster sampling is illustrated by the table below. The table gives population parameters.

Parameter symbol		Description	Central province Population	Central province Sample
Population	Sample			
N	n	Number of clusters	7 Districts	3 Districts
M_i	m_i	Number of elements	Kiambu = 853,772 Kirinyaga = 525,069 Muranga = 401,273 Nyandarua= 550,683 Nyeri = 755,588 Thika = 738,512 Maragua = 443,961	Kiambu= 59 Nyeri = 156 Kiambu = 153

Table 2 Parameter Symbols

Data is analysed using R and SPSS statistical software. Using the software R, a two stage cluster design is build which forms the basis of most of the analysis.

3.1 Descriptive statistics

3.2.1Estimation of the population totals

Taking into consideration that simple random sampling is adopted in the second sampling stage, the estimator of the population total would be:

$$\widehat{Y_i} = M_i\overline{y_i}$$

The unbiased estimator of $\hat{Y}$ would then be $\hat{Y} = \sum_{i=1}^{n} \frac{N_i}{n_i} Y_i$

The sampling variance would then be

$$V(\hat{Y}) = M^2(1 - f_1)\frac{S_1^2}{m} + \frac{M}{m}\sum_{i=1}^{N} N_i^2(1 - f_{2i})\frac{S_{2i}^2}{n_i}$$

3.1.2Estimation of means and proportions

3.1.2.1 Means

Estimate for the mean is given by the equation

$$\overline{\overline{y}} = \sum_{i=1}^{n} \frac{\overline{y_i}}{n} \quad \text{Where} \quad \overline{y_i} = \frac{y_i}{n_i} \text{ and } y_i = \sum_{j=i}^{m_i}(y_{ij})$$

The standard error of the mean is given by

$$\text{Se}\,(\overline{\overline{y}}) = \left[(1 - f_1)\frac{s_1^2}{n} + (1 - f_2)\frac{s_2^2}{mn}\right]^{\frac{1}{2}}$$

Where;

1. $s_1^2 = \frac{\sum_{i=1}^{n}(y_i - \overline{\overline{y}})^2}{n-1}$ is sample variance between total values of the characteristic Y per cluster *(M)*
2. $s_{2i}^2 = \frac{\sum_{i=1}^{mi}(y_{ij} - \overline{y}_i)^2}{m_i - 1}$ is Sample variance between values of the characteristic Y in the elements *(m)* sampled within cluster *(M_i)*
3. $s_2^2 = \frac{\sum_{i=1}^{n} s_{2i}^2}{n}$ is sample variance between values of the characteristic Y in the elements *(SU2)* sampled within all clusters *(M)*.
4. $f_1 = \frac{n}{N}$, $f_{2i} = \frac{m_i}{M_i}$, $f_2 = \frac{m}{M}$

3.1.2.2 Proportions

In the case of two-stage sampling, estimate the proportion of elements of the population belonging to one certain category can be calculated as estimated below.

A proportion in a sample of size n is considered as a mean of n Bernoulli variables.

Then, the proportion p_i, in the i^{th} sampled cluster i^{th} M, is: $p_i = \overline{y}_i$

Then the sample mean per element is the overall proportion:

$$\overline{\overline{y}} = \frac{\sum_{i=1}^{n} \overline{y}_i}{n} = \frac{\sum_{i=1}^{n} p_i}{n} = \overline{p}$$

Therefore, the estimator of the overall proportion of the elements belonging to the category of interest can be the average, $\widehat{P}$ of the proportions, p_i, of the clusters sampled: $\hat{P} = \overline{\overline{Y}} = \overline{p}$.

This estimate is unbiased and the sampling variance is given by;

$$v(\overline{p}) = (1 - f_1)\frac{s_1^{*2}}{n} + f_1(1 - f_2)\frac{s_2^2}{nm}$$

Where;

$$s_1^{*2} = \frac{\sum_{i=1}^{n}(p_i - \overline{p})^2}{n-1} \; and \; s_2^2 = \frac{\sum_{i=1}^{n} s_{2i}^2}{n} \; with \; s_{2i}^2 = \frac{mp_iq_i}{m-1}$$

3.1.3 Ratio estimation

In this study, one would be interested in analyzing the ratio of amount of money spent on foods, in restaurants, education, clothes, hospital, mortgage and rent etc, out of the total household expenditure. Also given the total number of hours in a week, one would be interested in the ratio

of time spent on paid work, leisure activity, collecting water, etc. this requires ratio estimation in complex survey methods.

Ratio estimate = $r = \frac{\mu_y}{\mu_x} = \frac{\mu_y expenditure\ on\ i}{\mu_x total\ expenditure}$, where *i* represent various household expenditures.

μ_y and μ_x could be represented by the characteristic mean or total.

This estimate also has it standard error.

3.2Test of association

Test of association is used to test if two variables are associated before a logistic regression model is fitted into the data. This is a relationship between two variables that renders them statistically dependent. Association refers to any such relationship where as correlation refers to linear relationship.

In order to test whether there is an association between two categorical variables, we calculate the number of individuals we would get in each cell of the contingency table. The contingency tables created by the described design care for the sampling weights. This gives an estimate of the province frequencies for various variables from the observed three districts.

3.3 Ordinal logistic regression (cumulative logit)

3.3.1 The Model

Ordinal response questions are common in survey practice. Many survey analysts do not hesitate to fit a standard linear regression model to ordinal response data. Others ignore the natural ordering of the response categories and analyse the ordinal data as though there was no order.

Analysts who ignore ordinality of such responses and apply the general multinomial logit regression technique of are also certainly not wrong in their approach. However, such models require the estimation of many parameters $(K - 1) \times (p + 1)$ parameters estimates are required when fitting a multinomial logit model and therefore may not be the most efficient modeling option.

Like most regression models, the interpretation of results from a cumulative logit regression model can occur at two different levels. At the evaluation stage, *t*-tests of single-parameter predictors or Wald tests of multiparameter predictors will identify those predictors that have a significant relationship with the ordinal response variable. Examination of the estimated coefficients for the cumulative logits can inform the analyst about the directional nature of the relationship of response and predictors.

A **cumulative logit** is defined for the probability of having an ordinal response less than or equal to *k*, relative to the probability of having a response greater than *k*: For an ordinal variable with *K* categories, $K - 1$ cumulative logit functions are defined. Each cumulative logit function includes a unique intercept or "cutpoint," $B_{0\,(k)}$, but all share a common set of regression parameters for the *p* predictors, $\boldsymbol{B} = (B_1 \ldots Bp)$. Consequently, a cumulative logit model for an ordinal response variable with *K* categories and $j = 1 \ldots p$ predictors requires the estimation of $(K - 1) + p$ parameters—far fewer than the $(K - 1) \times (p + 1)$ parameters for a multinomial logit model. The version that shows what function of the probabilities results in a linear combination of parameters is

$$In\left[\frac{prob(event)}{1 - prob(event)}\right] = \beta_o + \beta_1 x_1 + \beta_1 x_1 + \cdots + \beta_k x_k$$

The quantity to the left of the equal sign is called a **logit**. It's the log of the odds that an event occurs. (The odds that an event occurs are the ratio of the number of people who experience the

event to the number of people who do not. This is what you get when you divide the probability that the event occurs by the probability that the event does not occur, since both probabilities have the same denominator and it cancels, leaving the number of events divided by the number of non-events). The coefficients in the logistic regression model tell you how much the logit changes based on the values of the predictor variables.

For ordinal categorical variables, the drawback of the multinomial regression model is that the ordering of the categories is ignored.

Before beginning the model-building process, we examine simple frequency distributions for ordinal response variables. If the majority of responses on a discrete ordinal outcome are grouped in single categories or highly skewed to the highest or lowest possible values, the cumulative logit model may not be the best choice and simple or multinomial logit regression models would be more appropriate for a recoded version of the ordinal response variable.

In proportional odds model, the assumption is that the odds ratio assessing the effect of a predictor is the same regardless of where the cut point is made. In a 4 categorical response variable, the ratio that compares categories greater than or equal to 1 to less than 1 is the same as the odds ratio that compares categories greater than or equal to 3 to less than 3. In other words the odds ratio is invariant.

The form of proportional odds model with an outcome (D) with G levels (D=0, 1, 2… G-1) and one independent variable (X_1) is given by;

$$P(D \geq g|X_1) = \frac{1}{1 + \exp\left[-(\alpha_g + \beta_1 x_1)\right]}$$

Where g= 1, 2… G-1

$$1 - P(D \geq g|X_1) = 1 - \frac{1}{1+\exp\left[-(\alpha_g+\beta_1 x_1)\right]}$$

$$= \frac{\exp\left[-(\alpha_g+\beta_1 x_1)\right.}{1+\exp\left[-(\alpha_g+\beta_1 x_1)\right]}$$

$$= P(D < g|X_1)$$

$$Odds = \frac{P(D \geq g|X_1)}{1-P(D \geq g|X_1)} = \frac{P(D \geq g|X_1)}{P(D<g|X_1)}$$

$$= \frac{\frac{1}{1+\exp\left[-(\alpha_g+\beta_1 x_1)\right]}}{\frac{\exp\left[-(\alpha_g+\beta_1 x_1)\right.}{1+\exp\left[-(\alpha_g+\beta_1 x_1)\right]}}$$

$$= \exp\left[(\alpha_g + \beta_1 x_1\right]$$

1.3.2 Odds ratio and confidence limits

After the proportional odds model is fit and the parameters estimated, the process for computing the odds ratio is the same as in standard logistic regression. Consider a case where there is an independent variable with two levels. $X_1=\begin{cases}1 \text{ when in the category} \\ 0, \quad \text{otherwise}\end{cases}$

For comparing D

$$odds(D \geq g) = \frac{P(D \geq g|X_1)}{P(D < g|X_1)} = \exp\left(\alpha_g + \beta_1 x_1\right)$$

$$\text{Odds ratio} = \frac{P(D \geq g|X_1=1)/P(D<g|X_1=1)}{P(D \geq g|X_1=0)/P(D<g|X_1=0)}$$

$$= \frac{\exp\left[\alpha_g+\beta_1(1)\right]}{\exp\left[\alpha_g+\beta_1(0)\right]}$$

$$= \frac{\exp\left(\alpha_g+\beta_1\right)}{\exp\left(\alpha_g\right)}$$

$$= e^{\beta_1}$$

Confidence interval estimation is also analogous to standard logistic regression. The general large-sample formula for a 95% confidence interval, for any two levels of the independent variable (X_1 ** and X_1 *),

$$95\%\text{CI} = \exp\left[\hat{\beta}_1(X_1 ** - X_1 *) \pm 1.96(X_1 ** - X_1 *)s_{\hat{\beta}_1}\right]$$

CHAPTER IV

DATA ANALYSIS

4.1 Descriptive statistics

4.1.2 The design object

Data is analysed using both SPSS and R software's of data analysis.

A two stage cluster sampling object is created in R which form the basis of most of the analysis performed. The object describe the sampling design used in this case first stage sampling of the primary sampling units (PSU) and second stage sampling of secondary sampling units (SSU)

In the design the PSU's are the districts and the SSU's are the households selected from the selected districts.

The districts are selected with equal probability of selection. Districts are selected with probability 3/7. The households are selected with unequal probabilities.

$$Kiambu = {}^{59}/_{216693}$$

$$Muranga = {}^{153}/_{106686}$$

$$Nyeri = {}^{156}/_{205322}$$

To confirm the design we single out the variable that indicates the respondent's financial control in the household. The two tables show the outcomes.

How much control do you have in household decisions?

WITHOUT THE 2 STAGE CLUSTER DESIGN

		district			
		MURANGA	KIAMBU	NYERI	Total
	some control	22	9	26	57
	equal control	69	22	53	144
	most of the control	32	16	49	97
	complete control	30	12	28	70
Total		153	59	156	368

Table 3 Totals by variable "CONTROL" without the design

WITH THE DESIGN

Control	District			Total
	MURANGA	KIAMBU	NYERI	
some control	35794.43	77128.02	79847.44	192769.89
equal control	112264.35	188535.15	162765.94	463565.44
most of the control	52064.63	137116.47	150481.72	339662.82
complete control	48810.59	102837.36	85989.56	237637.51
Total	248934.00	505617.00	479084.66	1233635.66

Table 4 Totals by variable "CONTROL" with the design

The non- design based table 3 above gives the values for total number of respondents for the various category of the variable labeled control as observed from the three districts.

The design based table 4 gives the estimated total for the central district from the values collected from the three districts. The sampling probabilities are taken into account.

E.g. those who had some control in the whole province based on Muranga district were 35794 which is given by 22 *7/3*106686/153 where 22 is the observed count, 7/3 finite population collection based on PSU and 106686/153 is the finite population collection based on the SSU. The other values are determined likewise but the finite population collection changes from district to district.

The total number of households in the province based on the model object created and estimated from the variable labeled "CONTROL" is given in the table below.

	Total	SE
Some control	192770	37395
Equal control	463565	56635
Most of the control	339663	73648
Complete control	237638	41676
TOTAL	1233636	209354

Table 5 CONTROL totals

This shows that the total number of households in the district is between 1024282 and 1442990. The projected number of households according to Kenya national Bureau of Statistics was 1106909 which is between the values above. This confirms the design object created in R for analysis as a correct estimate.

4.1.2 Means proportions and totals.

The data has four variables that can be used as dependant variables.

a. Variable name (Control)

   The respondents were asked how much control they had in household decision making.

The response to this question was;

	Total	SE		Proportion	SE
some control	192770	37395	some control	0.15626	0.0159
equal control	463565	56635	equal control	0.37577	0.0279
most of the control	339663	73648	most of the control	0.27533	0.0265
complete control	237638	41676	complete control	0.19263	0.0179

Table 6 CONTROL by totals, proportions and the standard errors

b. Variable name (managMon)

The respondents were asked which member was responsible for day to day financial management.

The response to this question was;

	total	SE		Proportion	SE
No response	219917	57518	No response	0.178268	0.0434
Husband	395485	58725	Husband	0.320585	0.0208
Wife	493021	90857	Wife	0.399649	0.0409
Husband and wife together	64372	22743	Husband and wife together	0.052181	0.0130
Husband and wife separately	60841	21111	Husband and wife separately	0.049318	0.0131

Table 7 manaMon by totals, proportions and the standard errors

c. Variable name (finalSay)

The respondents were asked who the final decision maker was in the house.

The response to this question was;

	Total	SE		Proportion	SE
No response	290954.7	78121.7	No response	0.2358514	0.0543
Husband	549334.9	134911.9	Husband	0.4452975	0.0893
Wife	385393.9	86784.7	Wife	0.3124050	0.0435
Don't know/can't say	7952.1	4222.6	Don't know/can't say	0.0064461	0.0041

Table 8 finalSay by totals, proportions and the standard errors

d. Variable name (at-final)

The respondents were should make important decisions in the household.

The response to this question was;

	total	SE		Proportion	SE
No Response	152569	44909	No Response	0.12367	0.0356
Husband	235010	45512	Husband	0.19050	0.0438
Wife	846057	167440	Wife	0.68582	0.0369

Table 9 by totals, proportions and the standard errors

4.1.3 Ratio estimation

Some of the variables in the data were ratios of others. Household expenditure on various item were a ratio of the total household expenditure.

Ratios of population means or totals are important for three reasons;

1. the ratio may be of direct interest,
2. it may be used to estimate a population mean or total,
3. or it may be used to construct a subpopulation estimate of mean.

Estimation of the ratio of the amount of money spent on various household items by total household expenditure.

	RATIO	S.E
SP_FOOD	0.3120	0.01216
SP_EDUC	0.1339	0.02081
SP_VEGET	0.0883	0.00365
SP_HOUSI	0.0716	0.00571
SP_CEREA	0.0710	0.00266
SP_ENERG	0.0673	0.00691
SP_TRANS	0.0640	0.00753
SP_CLOTH	0.0545	0.00366
SP_DAIRY	0.0518	0.00307
SP_GRVEG	0.0370	0.00261
SP_PULSE	0.0332	0.00207
SP_GAS	0.0317	0.00195
SP_MEDIC	0.0296	0.00544
SP_PHONE	0.0267	0.00528
SP_OIL	0.0260	0.00123
SP_SNACK	0.0217	0.00109
SP_COFEE	0.0144	0.00177
SP_INSTA	0.0014	0.00070

Table 10 Proportion of household expenses

The table above shows that in central province, all population households spent 7.1% on cereals (SP_CEREA) with a standard error of 0.26567%, 3.695% was spent on green vegetables (SP_GRVEG), etc. The table shows the ratios of expenditure arranged in ascending order. Thus food takes the highest share followed by education, Vegetables, housing, in that order.

4.2 Exploratory data analysis

We look at the graphical nature of the data

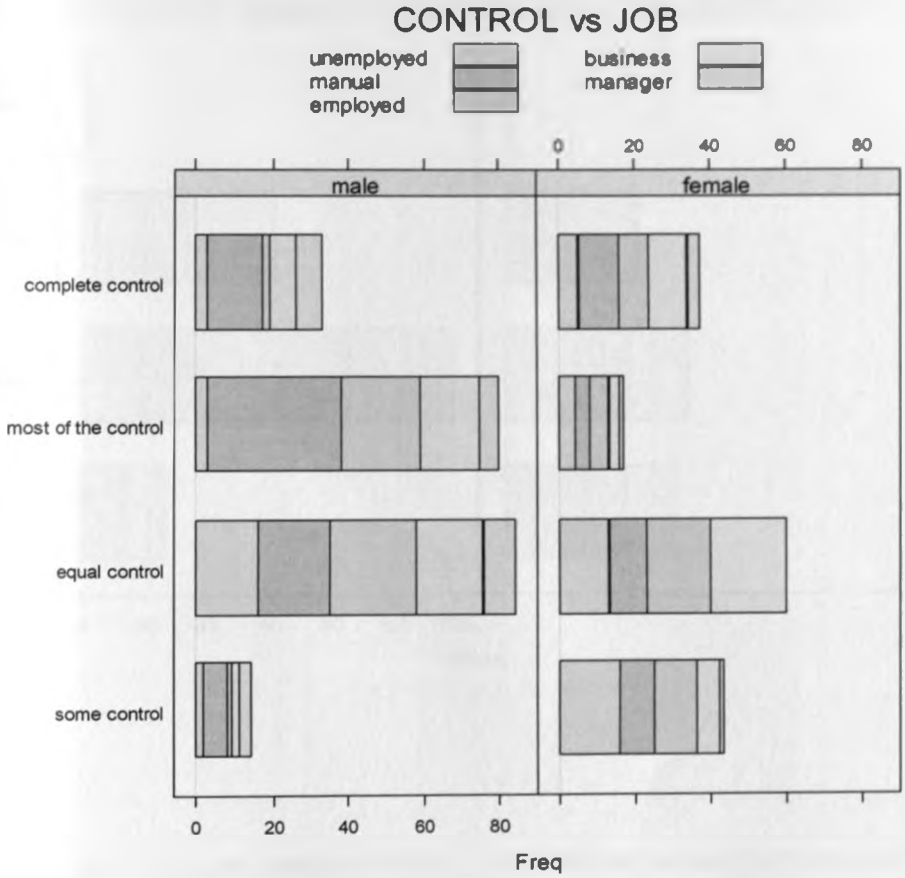

Fig 4.2.1

From the graph above, the male respondents who perform manual work had complete control of the household decision making, while unemployed female respondents had very little say. Most of the employed respondents had equal control. More males than females seem to have higher levels of control taking a cut from 'most of the control' to 'complete control'. In the higher levels of control majority of males are on manual jobs.

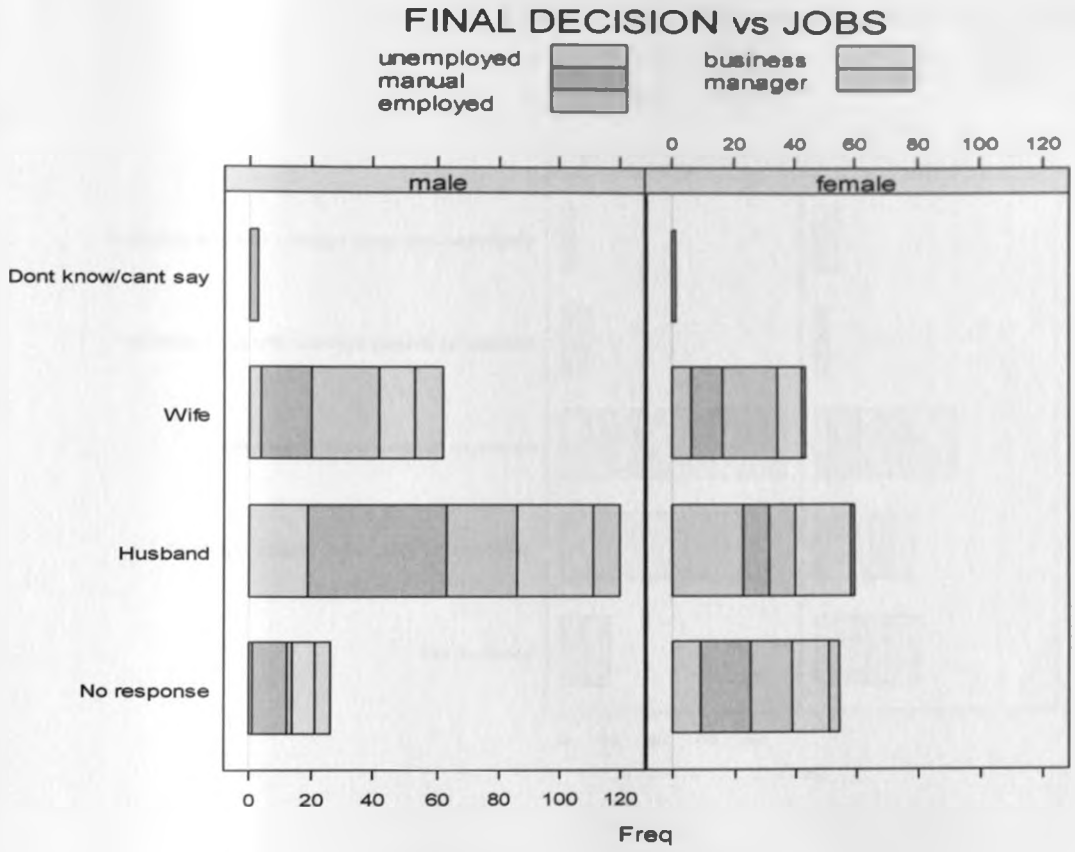

Fig 4.2.2

In the above figure females who are unemployed and males who do manual work were of the opinion that the husband makes the final decision, majority of the employed male and female respondents felt the wife should be the final decision maker in the household. More males than females felt that the husband should be the final decision maker. In this category a higher proportion was for males who were on manual jobs. More unemployed females than males felt that the husband should be the final decision maker.

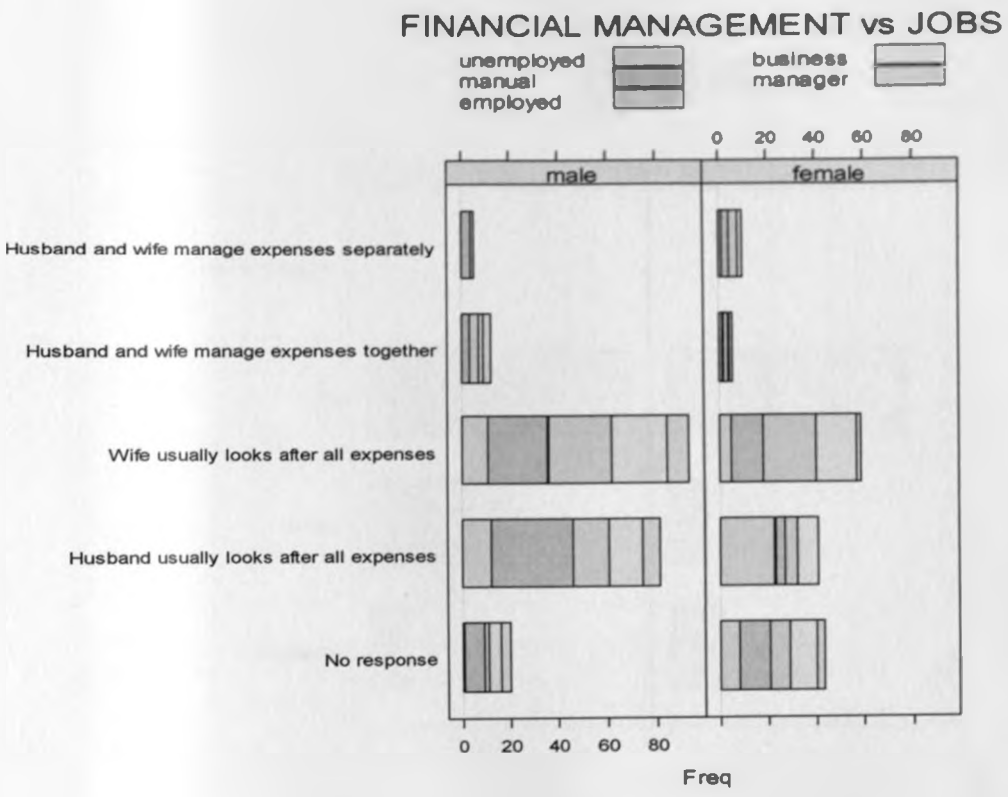

Fig 4.2.3

Majority of the households have the wife being responsible for day to day financial management. The majority of males on manual jobs have the husband being responsible for day to day financial management. Most of the females who are unemployed have the husband being responsible for daily financial management. In central province sharing of financial management is not practiced.

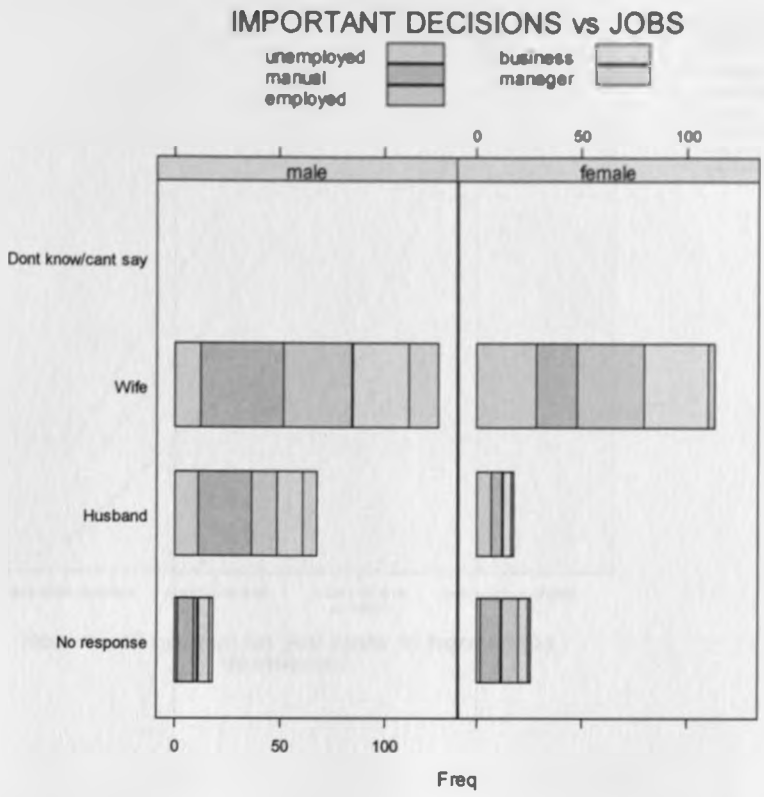

Fig 4.2.4

Clearly in central Kenya majority of the households important decisions are made by the wife but a big proportion of males in manual jobs has the husband making important decision in the household.

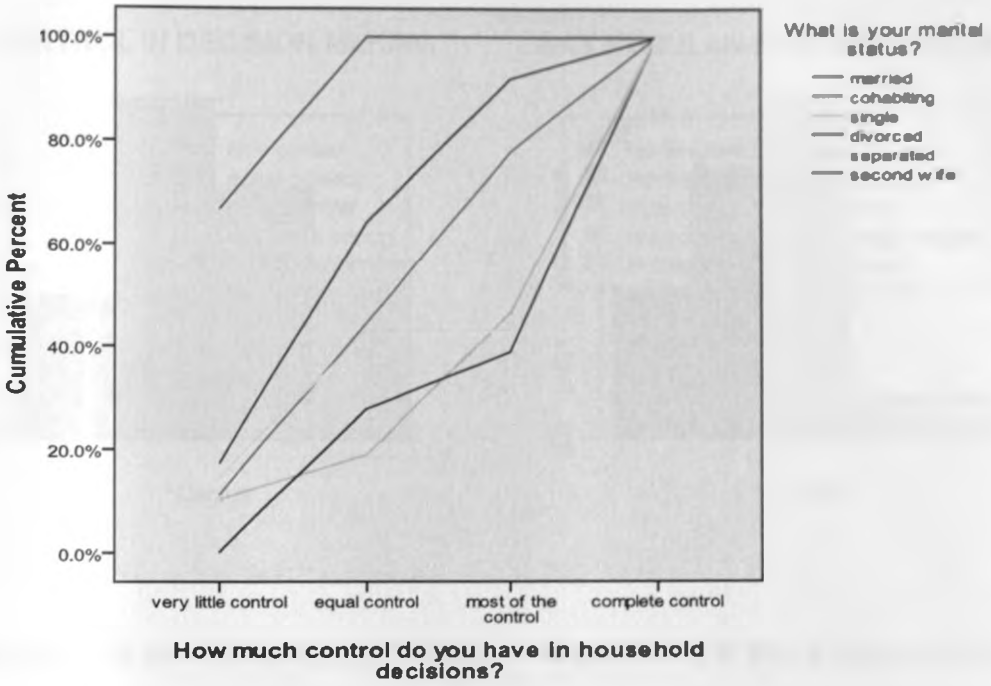

Fig 4.2.5

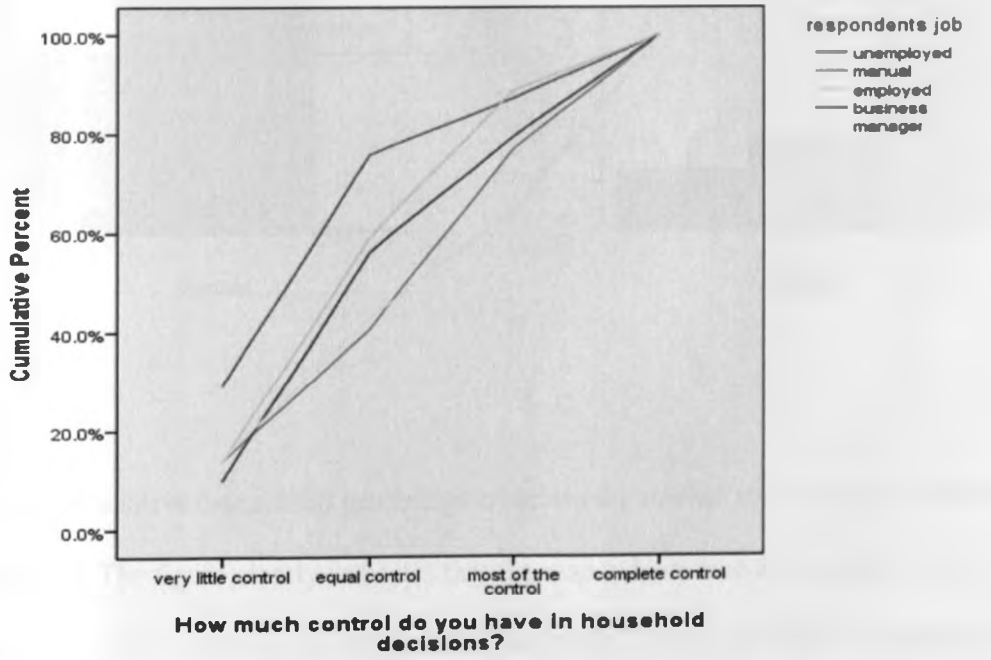

Fig 4.2.6

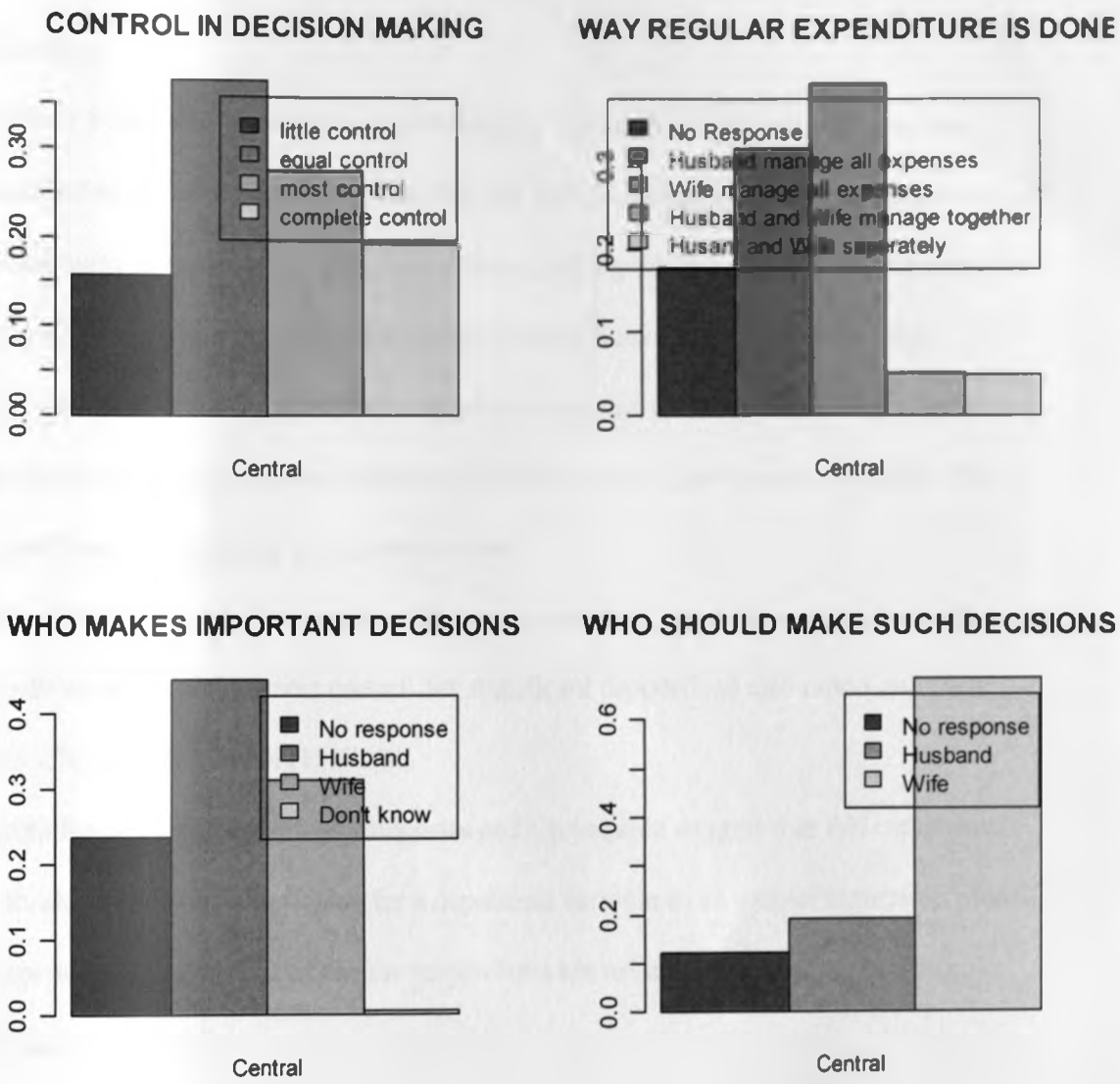

Fig 4.2.7

Fig 4.2.5 show a cumulative categorical percentage count on the marital status against household control of decisions. The figure clearly indicates that the respondents who are second wife are associated with very little control of decisions in the household. These are closely followed by those who are married and cohabiting. On the higher side of the graph there is an indication that

those who are single, divorced and separated will be associated with higher levels of household decision making.

Fig 4.2.6 show a cumulative categorical percentage count on the respondent's job against household control of decisions. Again it can be said that the unemployed and the employed will be associated with lower level of control of decision making while those who do manual work and the managers will be associated with higher levels of household decision making.

Fig 4.2 7 shows graphs for the identified categorical response variable. Both the graphs for control in decision making and way regular expenditure is done show some normality. This forms a good basis for applying a regression model.

We see from this figure that the majority of the respondents in population are estimated to respond with equal control to most control but significant proportions also report complete control (19.2%) or little control (15.6%).

This distribution of responses across categories and the implicit suggest that this categorical variable would be a reasonable choice for a dependent variable in an ordinal regression model. A similar experience is suggested where the respondents are asked how regular expenditure is shared in their households.

4.3 Test of association

Response variable	Predictor variable	χ-squared	Df	P- value
CONTROL	sex	74.48	3	2.184e-08
CONTROL	married	147.17	15	2.2e-16
CONTROL	Job3	63.64	12	2.296e-05
CONTROL	violFrie	24.52	3	0.00065
CONTROL	Viol`1_hit	18.43	3	0.01243
CONTROL	At_roles	38.83	12	0.01
CONTROL	Out_hus	53.87	21	0.01
CONTROL	At_final	53.85	6	6.277e-05
CONTROL	Final	168.61	9	2.2e-16
CONTROL	spouseIn	93.32	3	4.502e-14

Table11 Test of association

The associations show a strong association between the dependent variable Control and the independent variables sex, marital status, job, violence, role of the husband, spouses living together and drinking habits.

4.4 Cumulative logistic regression model

The results of the regression model fitted on the respondents control in household decisions (CONTROL) as the response variable and both job of the respondent (JOB) and the marital status of the respondent as the predictor variable.

TWO STAGE CLUSTER DESIGN-BASED RESULTS

PREDICTOR VARIABLE	CATEGORY	Value	Std.Error	t.value	p.values	Exp (β)
JOB	Manual	2.0058	0.6831	2.9364	0.0017	7.432037
	Employed	1.4673	0.5415	2.7096	0.0034	4.337508
	Business	1.7699	0.7463	2.3715	0.0089	5.870266
	Manager	1.9470	0.6781	2.8712	0.0020	7.007633
MARRIED	Cohabiting	1.0482	0.5371	1.9516	0.0255	2.852512
	Single	2.3255	0.4189	5.5512	0.0000	10.23179
	Divorced	3.6661	1.1483	3.1928	0.0007	39.09912
	Separated	4.1109	2.0952	1.9621	0.0249	61.00159
	second wife	-3.2257	1.5860	-2.0339	0.0210	25.17119(1)
SEX	Female	-1.4607	0.2807	-5.2032	0.0000	4.308974
RESPONSE VARIABLE		(intercept)				
little control|equal control		-0.0605	0.5280	-0.1146	0.4544	1.062368
equal control|most of the control		2.1490	0.6603	3.2546	0.0006	8.576278
most of the control|complete control		4.0096	0.8456	4.7418	0.0000	55.12482

(1) Indicate association on the lower side of the response variable

Table 12A Parameter estimates

CLASSICAL RESULTS

PREDICTOR VARIABLE	CATEGORY	Value	Std..Error	t.value	p.values	Exp (β)
JOB	manual	1.1178	0.3228	3.4631	0.0003	3.058119
	employed	0.8874	0.3255	2.7264	0.0034	2.428807
	business	1.0347	0.3314	3.1223	0.0010	2.814262
	manager	0.9248	0.4773	1.9377	0.0267	2.521364
MARRIED	cohabiting	0.8635	0.6534	1.3217	0.0936	2.371446
	single	2.6772	0.3385	7.9082	0.0000	14.54431
	divorced	3.2975	0.5493	6.0035	0.0000	27.04494
	separated	3.1471	0.9064	3.4722	0.0003	23.26849
	second wife	-2.2732	1.2363	-1.8386	0.0334	9.710425(1)
SEX	female	-1.3713	0.2358	-5.8168	0.0000	3.94047(1)
RESPONSE VARIABLE						
some control|equal control		-1.3054	0.2981	-4.3787	0.0000	3.689164
equal control|most of the control		1.0712	0.2882	3.7172	0.0001	2.91888
most of the control|complete control		2.7350	0.3210	8.5203	0.0000	15.40974

(1) Indicate association on the lower side of the response variable

Table 12B Parameter estimates

Table 12A results are derived from design-based model in which two stage cluster sampling weights are factored in where as Table 12B are derived from the classical methods in which the survey procedures are ignored. The β values differ in the two tables notably indicating the need to factor sampling weights when performing statistical data analysis of sampled data. The results in Table 12B would give misleading results.

To aid in the interpretation of the results the table 13 shows the various levels of the predictor variables used in the model above. Included in the table are the estimates of the totals for the whole of central province from the two stage cluster design created.

Comparing table 12 and table 13it is clear that one level in each of the predictor variables is missing. In the variable Job, the level unemployed is missing and this is the referent level. Likewise in the variable married, the level married is the referent level and for sex the referent is male. This combined with the exponent values helps us interpret the results.

job3	Respondents job (with five levels)					
		unemployed	Manual	employed	business	manager
		183576.56	403883.07	278141.78	268991.44	99042.82
Married	Marital status (with six levels)					
	married	cohabiting	Single	divorced	separated	second wife
	930162.29	27639.50	181123.17	65721.61	21219.97	7769.13
Sex	Sex of the respondent					
					male	Female
					720315.7	513320.0

Table 13 Totals by predictors

In Table 8A, the estimates of the four response variables dichotomized at four levels. These parameter estimates are rarely of real analytic interest but are used by the analyst to calculate predicted probabilities of being in one of the four ordered response categories for CONTROL.

Consider next the regression parameter estimates and estimated cumulative odds ratios (Table 12A) for the respondent job (Job) and marital status (married) predictors. Since each of these factors is represented by a single model parameter at each level, the *t*-test reported is equivalent to the Wald test of the significance of each predictor. The *t*-test results suggest that sex, job and marital status are significant predictors in the cumulative logit model for CONTROL. P- Values are also calculated to confirm these results. Positive β values are associated with higher levels of the response variable whereas the negative values are associated with lower levels of the response variable.

The value of exp(2.0058)= 7.432037 for the manual JOB indicate that those who perform manual jobs are 7.4 times more likely to have higher levels control of household decision making than lower levels of control to those who are unemployed. Proportional odds assumption states that the ratios assessing the effect of an exposure variable for any of these comparisons will be the same regardless the cut point is made. This implies that those who do manual jobs are also likely to have 7.4 times more control and most control to those who are unemployed. Those in this category are 7.4 times more likely to have higher levels of control than those who are unemployed. This agrees with the observations made on Fig 4.2.4. Those who are unemployed are placed the highest which means they are associated with little control as compared to those who perform manual jobs who are placed lowest meaning they are associated with higher levels of household decision making. Thus from table 12, it is can be seen that the employed are 4.34 time more likely to have complete control than these the unemployed. Those in Business are 5.87 more likely to have complete control than the unemployed and managers almost like those who do manual jobs are 7 times more likely to have complete control of household decisions. This

means if a seminar was held to address the concerns of complete household decision making e.g. how this would affect family relationships it would target the managers and the manual workers. On the marital status side of table 12, it can then be seen that those who are separated are 61 times more likely to have complete control of household decision. Those who are divorce are 39 times more likely to have complete control than the married, 10 times for the singles, 2.9 for those cohabiting but those who are second wives have 25.17 times of little household control than the married. The negative on the β means they are associated with the lower values. On gender the females are 4.3 times more likely to have little control than the males. It is worthwhile noting also that females are associated with lower levels of decision making.

Variable	Levels	Value	Std.Error	t.value	p.values	Exp (β)
violFrie	yes	0.6687	0.2089	3.2012	0.0007	1.9516
viol_hit	yes	0.7091	0.3428	2.0683	0.0193	2.0321
at_roles	Partly agree	0.2231	0.2923	0.7633	0.2226	n/a
	Neither	0.6872	0.4620	1.4876	0.0684	n/a
	Partly disagree	-0.0821	0.4626	-0.1775	0.4296	n/a
	Strongly disagree	-0.2404	0.3728	-0.6448	0.2595	n/a
CONTROL (intercept)	very little control|equal control	-1.1271	0.3214	-3.5065	0.0002	
	equal control|most of the control	1.0754	0.3996	2.6914	0.0036	
	most of the control|complete control	2.8331	0.4401	6.4376	0.0000	

Table 14 Parameter estimates

Viol-Frie is a variable that indicates whether any of the respondents close friends has been a victim of domestic violence. Vio_ hit indicates if the respondent has ever hit his/her partner. At_roles is a variable that indicate the respondent's opinion that it's the husband's role to look for money and the wife's role to look after home and family.

The results of the cumulative odds model show that only Viol_Frie and Vio_hit are significant. Results show that those who fight their spouses are 2.0321 more likely to have complete control

of household decision making than those who do not. Those whose friends have been victims of domestic violence are 1.95 more likely to have higher levels of control those who have not.

VARIABLE	Level	Value	Std.Error	t.value	p.values	Exp (β)
out_hus	Once	0.2617	0.2737	0.9563	0.1695	
	Twice	-0.4938	0.4104	-1.2033	0.1144	
	Thrice	0.3821	0.3428	1.1148	0.1325	
	four times	1.4112	0.9103	1.5502	0.0605	
	five times	0.2004	0.3442	0.5822	0.2802	
	six times	0.9648	0.1184	8.1495	0.0000	2.62426
	every day	1.3941	0.7874	1.7705	0.0383	4.03134
CONTROL	very little control|equal control	-1.7269	0.1879	-9.1915	0.0000	
	equal control|most of the control	0.2253	0.1184	1.9018	0.0286	
	most of the control|complete control	1.7043	0.1587	10.7398	0.0000	

Table 15 Parameter estimates

The variable (out_hus) indicates the number of times the husband goes out to bar/club per week. Those who go out six times and every day have significant high levels of control of household decisions. More specifically those who go out six times are 2.62426 more likely to have higher levels of control than those who don't go out and those who go out every day are 4.031 more likely than those who do not go out at all.

Variable	Level	Value	Std..Error	t.value	p.values	exp(β)
spouseIn	yes	-1.973	0.3373	-5.8519	0.000	7.19(1)
CONTROL	very little control|equal control	-3.340	0.3185	-10.485	0.000	
	equal control|most of the control	-1.3718	0.3507	-3.9114	0.000	
	most of the control|complete control	0.2090	0.2401	0.8707	0.192	

(1) Indicate association on the lower side of the response variable

Table 16 Parameter Estimates

Table 16 indicates that the respondents who live with their spouse have 7.19 less more likely to have lower levels of control than those who live apart.

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions of the study

It is important to appreciate the use sampling techniques in research to reduce costs for studies and increase precision of measurements. In this study only 368 out of 1106909 households were included for the study but the results given were true projected estimates of what would have been achieved if all individual households were selected. It's important to include weights when analysing a design based data.

Majority of the respondents in central province Kenya felt that the husband should be the final household decision maker (45%±9%). Those who felt that the wife should be the final decision maker were fewer (31%±4%). It is evident from the results that the sharing of financial management is not a common practice in central province. Either the wife is responsible for day to day management (40% ±5% response) or the husband is (32%±2%). Only 5%±1% share the responsibility together and the same percentage share the responsibility separately. On who should make important decisions 19%±4% felt the husband should and 68% ±4% felt the wife should. This shows that although majority of wives make important decisions in the household, only a few have the final decision to those decisions.

The respondents were asked how much control they had in household decision making and this was seen to be strongly associated with the sex of the respondent, marital status, the nature of his/her job, social attitudes and whether the spouses live together. To address the concerns of high levels of household control which may affect family relations and progress, those who perform manual jobs and managers should the first target of sociologist and marriage experts

among the employment sectors. Those who do manual jobs due to the pressure they have at their work places compensate by being at high levels of control at their households. In the spirit of gender equality it is important that central province families engage equally and more so the husband involve their wives in decision making.

The next target would be the second wife who seems to have no say at all in the decision making. Respondents who are separated indicated the highest level of control than those who are divorced who were followed by the singles. Polygamous marriages thus should be discouraged for the sake of harmony in marriages. The two types of existing marriages that were studied are the cohabitation and legal marriages. Cohabitation showed a higher level of household decision making than the legal marriages.

Those who drink daily or six times a week are associated with high levels of household control. This indicates that alcohol can influence household decision making.

Violence in household is also a factor that can affect decision making by creating fear among spouses and negatively imparting on equality of making household decisions. This is accelerated by the fact that even those whose close friends who fight their spouses also have high levels of decision making. As the old adage go, birds of a feather flock together.

It is worthwhile noting that;

1. A decision may be reached by one person giving up his or her position to reduce or settle the conflict.

This may be an easy way out for less vocal spouse, but there is a risk of cutting off communication with the spouse and increasing stress. This is indicated by those whole have low levels of decision making. Those who are second wives belong to this category.

2. Another style is when spouses take an active role in the decision-making process by each person taking a turn "stating their case." Everyone then has a say in the outcome. Spouses who use this style are typically more flexible and open because they seek input from each member. This is indicated by those in the mid levels of decision making. Those who are legal marriages and those who are in cohabitation belong to this category.

3. Sometimes spouses have trouble getting past the brainstorming part of the process so one person may make a decision for the whole family. The spouses' reaction to this decision-making style can be either positive or negative. The other spouse may feel angry or hurt because they have no control or power over the decision, or they may be relieved that a decision has finally been made. This is indicated by those in the higher levels of decision making variable. Those who are single, separated and divorced belong to this category.

Applying classical statistical methods in analysis of survey data without making allowance for the survey design features can lead to erroneous inferences. In particular ignoring the survey design can lead to serious underestimation of parameter estimates as shown on tables 12 A and B and hence result to misleading results.

5.2 Recommendations

It is important that sociologist and marriage experts address nature of employment, type of marriage, social attitudes and income levels to be able to address the issues of equality in decision making for the benefits of household relationships.

5.3 Limitations of the study

when using the cumulative proportional odds models goodness of fit measures and diagnostic tools like those available for multivariate logistic regression have not been developed for the

complex logistic regression models for ordinal response data and therefore are not yet available in statistical software procedures for complex survey data analysis.

5.4 Areas of further research

Further research needs to be done do differentiate if the results would be the same if the control of household decision was categorized by area of control like education, foods, clothing, farming, etc. This is due to the fact that some of this decision depends on what is to be decided about.

Decision making can alter as the family ages. Studies need to be carried out to reveal how decision can be influenced at earlier family life, middle and later family life. This can be done by including a variable that records the number of years a couple have been married or have cohabited.

It is also necessary to determine whether there is any relation between sex differences and type of profession, and whether the age differences in the importance assigned to factors that affect decisions can be interpreted from a perspective of general experience -in the sense that as one gets older, one becomes more skilled at making decisions- or from the viewpoint of specific expertise, which suggests that adults and retired people are only more skilled in the domains in which they are more knowledgeable.

References

1. B. S. Everitt. A Skrendal 2010. The Cambridge Dictionary of Statistics 4th Edition. Cambridge university press.
2. Catherine Hobart Warner (2007). Mine is yours; Modes of expense sharing in married and cohabiting households.
3. David G. kleinbaum Mitchel Klein (2002). Logistic regression; A self learning text. Springer.
4. Emygdio Landerset Cadima, Ana Maria Caramelo, Manuel Afonso-Dias, Pedro Conte de Barros, Merete O. Tandstad (2005). Sampling methods applied to fisheries science; A manual. Food and agriculture organization of the united nations.
5. John M. Quick (2010). Statistical analysis with R. Pack Publishing.
6. John Simister (2010). Household financial management; evidence from 'work, attitude and spending' surveys. Was Research Paper.
7. Maria L, San de Acedo lizarraga, Maria T. Sanza de Acedo Baquedo, Ymaria Cardelle-elawar (2007). Factors that affect decision making gender and age difference. International journal of psychology and psychological theory.
8. Mehdi Yadollahi, Laily Hj Paim, Mumtazah Othman, Turiman Saundi (2009). Factors affecting family economic status European journal of scientific research. Page94-109.
9. N. H. Bingham, John M. Fry (2010). Regression; Linear model in statistics. Springer
10. Nora F. Almosead (2008). Money and power in saudi family. JKAU: arts and humanities page 61-87.
11. Peter Dalgaad (2008). Introductory to statistics with R. springer.
12. Phil Spector (2008). Data manipulation with R. Springer.

13. Risto lehtonen and Erkki Pahkinen (2004). Pratical Methods for Design and Analysis of Complex Surveys. John Wiley and sons, inc.

14. Rudolf J. Freund, William J. Wilson and Ping Sa (2009). Regression analysis; Statistical modeling of a response variable. Academic Press

15. Steven G. Heering, Brandy T. West, Patricia A Berglund (2010). Applied survey data analysis. CRC Press Taylor and Francis Group.

16. Thomas lumeley (1969). Complex survey: a guide to analysis using R. John Wiley and sons, inc.

17. William G. Cochran (1977). Sampling techniques 3rd edition. John Wiley and sons. inc.

18. Wyne A. Fuller (2009). Sampling statistics 2nd edition. John Wiley and sons, inc.