International Journal of Mathematics and Physical Sciences Research ISSN 2348-5736 (Online)
Vol. 4, Issue 2, pp: (64-84), Month: October 2016 - March 2017, Available at: www.researchpublish.com

Linear Discriminant Modeling of Wage Nonfarm Employment between Women and Men in Rwanda

[1]Anatole Mulindwa, [2]Joseph K. Mungatu Phd, [3]Marcel Ndengo Phd,

[1, 2,3] Affiliated to Jomo Kenyatta University of Agriculture and Technology (JKUAT). Kigali Campus, Rwanda

Abstract: **The ultimate goal of the study was to find out statistical linear model which predictor variables exactly discriminate or separate women and men groups in wage non-farm employment sector in Rwanda. Linear discriminant method was used with numerical data given by third EICV conducted from 2010-2011 published by NISR in 2012. Discriminant analysis assigns observations to one of the pre-defined groups based on the knowledge of the multi-attributes. For this study I had a single classification variable as sex (male and female) that were divided into two groups of male workers and female workers in non-farm works and the distribution with each group was multivariate normal. The research's sample was limited to the age between 18 and 65 years old by which women and men who are engaged in the wage nonfarm employment sector. This implied that 7,353 individuals belonged to the actual sample size with 3,772 (51.3%) women and 3,581 (48.7%) men. Majority of respondents were between 18 and 32 years old. 80% of the respondents had been to school and the level of Diploma is at 1%, Bachelor with 0.7%. 55% of the NFE workers are in the trade businesses. In the Non-Farm Employment sector. The SPSS was used to perform tests including the ANOVA test, test of variance, test of equality of group means, the Box's M test, the Wilks' Lambda test and Canonical discriminant analysis. Dependent variable was the sex type of male and female which was categorical and independent variables were: Type of non-farm activity (enterprise group), Education level, Income, Income-Unit of time, Expenditure, Expenditure-Unit of time, Duration, Urban/Rural location, and Poverty. The best predictors variables of the discrimination between women and men in the NFE sector were: Education level, Income, Income-Unit of time, Expenditure, Expenditure-Unit of time, Urban/Rural location, and Poverty and the weak predictor variables were: Duration in the business and Industry group of jobs.**

Keywords: **discrimination, women, men, nonfarm employment sector and statistical linear model.**

1. INTRODUCTION

1.1 Background of the study:

A non-farm household enterprise is an owner-operated business that is normally in the informal sector and does not employ paid workers on a regular basis. Other family members may contribute unpaid labor. 'Non-farm' refers to those activities that are not primary agriculture, forestry or fisheries but does include trade and the processing of agricultural products even when undertaken on the farm (Murenzi, Pamela, & Musana, 2012).

The issue of delivering long-term strong and sustainable economic growth that benefits *all* can only be met if best use is made of all available resources. Sendoff women behind means not only forsaking the important contributions women make to the economy but also wasting years of investment in education of girls and young women. Creation of the talent pool ensures that men and women have an equal chance to contribute both at home and in the workplace, thereby enhancing the well-being of both men and women, and more generally to society.

Who is a worker? An individual who works part-time or full-time under a contract of employment, whether oral or written, express or implied, and has recognized rights and duties

So what is a contract of employment? Oral or written, express or implied, agreement specifying terms and conditions under which a person consents to perform certain duties as directed and controlled by an employer in return for an agreed upon wage or salary. Whether stated or not in the contract, both the employee and the employer owe the duty of mutual confidence and trust, and to make only lawful and reasonable demands on each other. Every employee is under the obligation to carry assigned duties, or the employer's instructions to the best of his or her abilities (Moss, 2008).

Greater educational equality does not guarantee equality in labor market outcomes, however. If high childcare costs mean that it is economically not worthwhile for women to work full-time, if workplace culture penalizes women for taking a break to have a child or provide for elderly relatives and as long as women continue to bear the main brunt of unpaid household tasks, childcare and caring for ageing parents, it will be difficult for them to realize their full potential in paid work. In developing countries, if discriminatory social norms enhance early marriages or limit access to credit for women, the significant gains made in educational attainment for girls may not lead to increased formal employment and entrepreneurship (OECD, 2012).

But irrespective of family commitments, many female professionals find it difficult to climb the career ladder. In fact, inequalities increase the higher up the pay scale you go, so that while on average in OECD countries women earn 16% less than men, female top-earners are paid on average 21% less than their male counterparts. In Japan and the United Kingdom, the proportion of graduates with top grades is around 10 to 15% lower than for men, while in Estonia, Italy and the Netherlands it is the other way around. On average in the OECD in 2010, 65% of women were in the labor force up from 58% in 1990. However, there is considerable cross-national variation. In 2010, female labor force participation ranged from over 75% in China, the Nordic countries and Switzerland to below 50% in India, Mexico, South Africa and Turkey (OECD, 2012).

In fact, recent decades have seen a ‘feminization of agriculture’, in many developing countries. In 2010, 58% of women compared with 52% of men in Eastern and Middle Africa worked in the agricultural sector, whilst in South Asia these proportions were 51 and 35%, respectively. Women often remain marginalized in lower status, unskilled agricultural work, which is frequently unpaid (Jütting, Luci, & Morrison, 2010). In general, women tend to be over-represented as contributing family workers and under-represented as employers.

On average in the OECD area, while 79% of women with tertiary education were in paid work, only 48% of women with less than upper secondary education was employed. Similarly, for example, in Egypt and Jordan, the employment rates of women with post-secondary education are three times as high as for the female population in general. The increase in female labor force participation over the past two decades has been slight: from 22% in 1990 to 30% in 2010, almost 40 percentage points below male labor force participation rate in the region. There is considerable cross-national variation: female labor force participation rates range from 15% in Iraq to 53% in Qatar. In Morocco, only 12% of married women join the labor force (compared with 79% of married men) (OECD, 2012).

Despite the rapid process of urbanization observed in most developing and transition countries, poverty still remains a predominantly rural phenomenon. Not only do a majority of the poor live in rural areas, but also the severity of their destitution is, on average, far greater than in urban areas. These trends are expected to persist in the foreseeable future. That being the case, it is critical that rural poverty is addressed in both poverty reduction strategies and, generally, as part of policies seeking to promote rural development. Correspondingly, it is important for developing countries and international development organizations to continue to assess approaches to rural development and their effectiveness in reducing rural poverty (Carletto & Katia, 2007).

The probability of access to and the level of non-farm income are functions of incentives offered by the economic context, the capacity of households to respond to those incentives, which are dependent of the household assets, and access to credits and government transfer (HUNG, 2006). Non-farm employment opportunities in rural areas likewise depend on proximity to urban centers, as do agricultural incomes (UN, 2015).

1.1.1 Rwanda context of non-farm employment:

Rwanda has been recognized as one of the best performing countries in Africa and as having created a very favorable 'soft' business infrastructure. Lacking natural resources, it has identified building a knowledge based economy as the path it must take to sustained economic development. It has achieved political stability and is widely recognized as having a committed, credible and capable government. However, Rwanda remains one of the poorest countries in the world and is heavily aid dependent, with over 50 per cent (52.48) of revenues in 2009 coming from aid. It is resource-scarce and landlocked, both strong inhibitors to economic growth. It remains heavily dependent on rain-fed, mainly subsistence, agriculture and there is high pressure on land, much of which is degraded through exhaustion and erosion. Seventy per cent of the population is young adults and children and approximately 85 per cent of the population lives in rural areas. High population growth (2.8% per year) combined with many farms too small to support a family means that there is an urgent need to increase non-farm employment opportunities and grow the labor market (Murenzi et al., 2012).

The overall Rwandan economy is growing at a significant rate. The average annual growth rate in GDP was 8.8 per cent between 2005 and 2009. Rwanda's GDP per capita has increased from less than 200US$ in 1994 to 540 US$ in 2010.Although still at an early stage, the GoR has set a set path towards economic transformation which shows signs of economic transformation in Rwanda. There was an increase in wage employment in the same period, from 11 per cent to 21 per cent, driven by an increase in wage farm employment (from 4% to 8%) as well as non-farm waged employment (from 7% to 11%). There was also an increase in employment in independent non-farm enterprises (from 5% to 10%). Considering jobs started in the last five years (excluding independent farmers) we can see that just under half (45%) were as unpaid farm worker while 12 per cent were in paid farm work; thus over 40 per cent were in non-farm employment. Twenty-three per cent were in paid non-farm employment and 19 per cent in independent non-farm work (16% as operator and 3% as dependent worker). Independent non-farm work together with waged non-farm was especially important for young people aged 21 to 30 years, with just over fifty percent of them starting non-farm employment (Malunda, 2012b).

The Interim Demographic Health and Survey (IDHS) 2007/2008 highlighted employment options available to young adults: 34% were engaged as unpaid workers (e.g. field crop and vegetable farm workers), 28% as independent farmers (e.g. general farmers), 15%) as wage farmers (e.g. field crop and vegetable farm workers), 12% as wage non-farmers (e.g. maids and brick layers) and 10% as independent non-farm (e.g. sales workers or street vendors) (MoH, 2015).

In Rwanda, agriculture sector has played a significant role in poverty reduction between 2001 and 2011with non-farm wage employment (3%), non-farm self-employment (13%), decreased dependence ratio (9%), increased agricultural commercialization (10%), increased agricultural production (35%), and other factors and unexplained part (30%) (Roger, Dan, & John, 2014). Poverty fell in almost all categories, but particularly among those reliant on non-farm wage or self-employment work, or transfers (Malunda, 2012).

For non-farm workers there are three work type classifications: the paid employee or waged nonfarm worker; independent self-employed persons and owners of small businesses; and the unpaid non-farm worker who work for no pay or reward in family enterprise, but who benefit from their work as a member of the household owning the business. In EICV1 85% of working adults were classified as subsistence farmers, either as an independent farmer or as a member of a family working on the family farm. By EICV2 this proportion had declined to 71%, with growth in waged farm labor, waged non-farm work and independent small business self-employment and in all provinces there has been an increase in wage farm labor and in non-farm jobs (NISR, 2007). Non-farm jobs increased from 17% to 20% (NISR, 2011).

Women have lower earnings and economic opportunities than men and most of their occupations are low-paying. Women tend to work in the agriculture sector at higher rates than men (82 percent versus 61 percent, respectively). Over time, men have increasingly moved from agricultural employment to non-farm work including non-productive activities and marketing. However, women face fewer opportunities than men when it comes to moving into non-agricultural jobs; from the years 2001 to 2006 12.3 percent of men were able to move out of the agricultural sector compared to 6.1 percent of women (USAID/Rwanda, 2015).

Although the increase in waged farm jobs is the same number for both sexes, the biggest numerical change for men is in non-farm work (Mary, Yussuf, & Emily, 2011).

In Rwanda, most of people aged 16 years and above are independent farmers (57.4%). Among them 21.6% are employed in the wage non-farm works (NISR, 2012). Households reliant on non-farm self-employment and especially non-farm wage work tend to be much less poor (NISR, 2012). The majority of Rwandan workers are engaged in non-wage employment. Even so, the percentage of the labor force in non-wage employment fell from 73% in 2005/06 to 64% in 2010/11 (AfDB, 2014). Majority of youth are engaged in unpaid work (34%) independent farm (28%) wage farm (15%), wage non-farm (12%) and independent nonfarm 10% (Ministry of Youth, 2010). Independent non-farm business owners spend on average 36 hours a week on their business (GoR & One UN, 2013). The primary farmer allocates significantly less time to own-farm agriculture and more time to non-agricultural work (Kati, 2008). The largest concentration of vulnerable or poor people in rural areas is in farming activities and most are contributing family workers (Demeke, Guta, & Ferede, 2006).

1.1.2 Aspects of Multivariate Analysis:

The study aimed at measuring gender pay gap in Rwanda using linear discriminant analysis. The research question was that do gender differences exist in wage Non-Farm Employment sector in Rwanda? We intend to show up statistical model that separates different determinants of difference between female workers' group and male workers' group in wage Non-Farm Employment sector.

Scientific inquiry is an iterative learning process. Objectives pertaining to the explanation of a social or physical phenomenon must be specified and then tested by gathering and analyzing data. In turn, an analysis of the data gathered by experimentation or observation usually suggests a modified explanation of the phenomenon. Throughout this iterative learning process, variables are often added or deleted from the study. Thus, the complexities of most phenomena require an investigator to collect observations on many different variables.

Because the data include simultaneous measurements on many variables, this body of methodology is called multivariate analysis. Though for this research we deal with only two variables with fisher discrimination and classification.

The need to understand the relationships between many variables makes multivariate analysis an inherently difficult subject. Often, the human mind is overwhelmed by the sheer bulk of the data. Additionally, more mathematics is required to derive multivariate statistical techniques for making inferences than in a univariate setting. We have chosen to provide explanations based upon algebraic concepts and to avoid the derivations of statistical results that require the calculus of many variables.

Multivariate analysis is a "mixed bag." It is difficult to establish a classification scheme for multivariate techniques that both is widely accepted and indicates the appropriateness of the techniques. One classification distinguishes techniques designed to study interdependent relationships from those designed to study dependent relationships. Another classifies techniques according to the number of populations and the number of sets of variables being studied.

The objectives of scientific investigations to which multivariate methods most naturally lend themselves include the following:

1. *Data reduction or structural simplification.* The phenomenon being studied was represented as simply as possible without sacrificing valuable information.
2. *Sorting and grouping.* Groups of "similar" objects or variables were created, based upon measured characteristics. Alternatively, rules for classifying objects into well-defined groups were used.
3. *Investigation of the dependence among variables.* The nature of the relationships among variables was of interest. All the variables mutually independent were one or more dependent on the others
4. *Prediction.* Relationships between variables were determined for the purpose of predicting the values of one or more variables on the basis of observations on the other variables.
5. *Hypothesis construction and testing.* Specific statistical hypotheses, formulated in terms of the parameters of multivariate populations, were tested.

1.1.3 Linear Discriminant Analysis:

Linear Discriminant Analysis, or simply LDA, is a well-known classification technique that has been used successfully in many statistical pattern recognition problems. It was developed by Ronald Fisher, who was a professor of statistics at the University College London, and is sometimes called Fisher Discriminant Analysis (FDA). The primary purpose of LDA is to separate samples of distinct groups as follow:

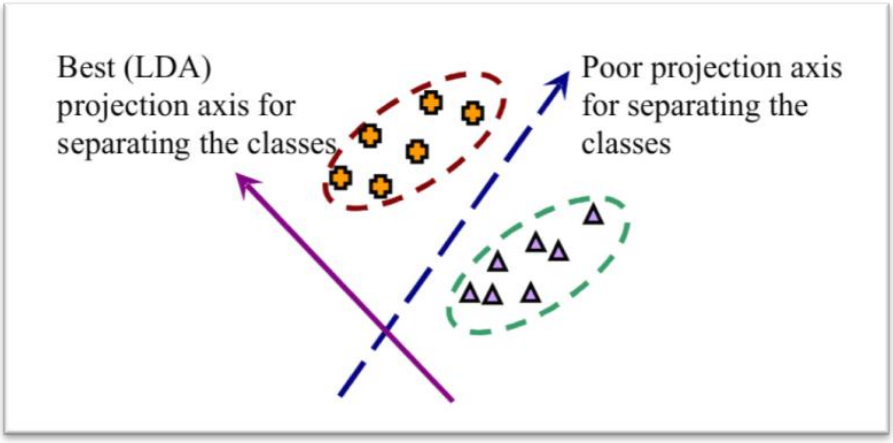

Figure 1: The projection axis for separating the classes

The LDA aims at finding two scatter matrices referred to as the "between classes" and "within class" scatter matrices. Suppose in a given problem we have g different classes or (sample groups). Each sample group πi has a class mean, which we denote $\bar{x}i$

What is Discriminant function analysis?

(i) It builds a predictive model for class membership

(ii) The model is composed of a discriminant function based on linear combinations of predictor variables.

(iii) Those predictor variables provide the best discrimination between classes.

Purpose of Discriminant analysis is:

(i) To maximally separate classes.

(ii) To determine the most parsimonious way to separate classes

(iii) To discard variables which are little related to class distinctions

Let us see some examples of where separation and classification for two populations were applied:

Table 1: Examples of separation and classification of two populations

N0	**Populations π_1 and π_2**	**Measured variables X**
01	Solvent and distressed property-liability insurance companies.	Total assets, cost of stocks and bonds, market value of stocks and bonds, loss expenses, surplus, amount of premiums written.
02	Nonulcer dyspeptics (those with upset stomach problems) and controls ("normal").	Measures of anxiety, dependence, guilt, perfectionism.
03	Males and females.	Anthropological measurements, like circumference and volume on ancient skulls.
04	Federalist Papers written by James Madison and those written by Alexander Hamilton.	Frequencies of different words and lengths of sentences.
05	Two species of chickweed.	Sepal and petal length, petal cleft depth, bract length, scarious tip length, pollen diameter.
06	Purchasers of a new product and laggards (those "slow" to purchase).	Education, income, family size, amount of previous brand switching.
07	Successful or unsuccessful (fail to graduate) college students.	Entrance examination scores, high school grade point average, number of high school activities.
08	Good and poor credit risks.	Income, age, number of credit cards, family size.
09	Alcoholics and nonalcoholic.	Activity of monoamine oxidase enzyme, activity of adenylate cyclase enzyme.

Source: Dickinson, D. L., Oaxaca, R. L., & Dickinson, D. L. (2006). Statistical Discrimination in Labor Markets: An Experimental Analysis, (2305).

If the values of X were not very different for objects in π_1 and π_2, there would be no problem; that is, the classes would be indistinguishable, and new objects could be assigned to either class indiscriminately.

1.2. Statement of the problem:

Gender equality is a multidimensional term embracing economic, cultural and social dimensions alike. Here, I confined myself to the important aspect that serves the purposes of my report. The equal right (and opportunity) to work and equal participation in the labor market. There are major differences in the levels of labor force participation between women and men. This is to a great extent due to perceptions about the role of women in the interaction between housework and work in the market. Traditionally, women have been expected to perform most of the work in the home as a matter of course, regardless of which partner is most suited to the task. This traditional attitude is still an important explanation of the differences in women's labor participation (considerably greater than men's) found in Rwanda especially in rural areas. The education level of women and men is essential condition of labor market equality though it is not sufficient itself.

Non-farm employment in Rwanda has long been seen by farm residents as a way to bridge the income gap among them that arises from stagnating farm production and growing population pressure. In Rwanda, where population density in certain regions approaches 428 persons per km^2 in 2011 and 460 persons per km^2 in 2014

However, women account for more than half of Rwanda's workers, but men are more likely to have wage non-farm employment. In fact, a large percentage of women work without pay (Rwanda, 2010). It also revealed that in wage non-farm employment sector in urban areas (21.6%) in rural areas (57.6%) and women are 8.3% while men are 27.4%. Female household heads are concentrated in agricultural jobs as the main usual occupation. 76% are independent farmers and a further 10% are waged farm workers, meaning 86% of female household heads work in farming. This compares with 62% of male heads (NISR, 2013).

There are almost 2 million female small-scale farm workers compared with just over 1.1 million men. it can be seen that the proportions of the workforce taking this kind of work fall dramatically in both sexes in the younger age groups; over 80% of women over 50 years of age do family farm work compared with just 60% of women under 25 year olds. For men the proportions decrease from over 60% in family farming among those over 50 years of age to around 43% for those under 25 (NISR, 2011).

Men worked seven hours more than women on average, but this is more than compensated for by the additional domestic duties undertaken by women. Summing domestic and working activities, men spent 40 hours working per week and women 51 hours a week. On average, men spent nine hours per week on domestic duties, while women spent 27 hours. Of all the duties, foraging for fodder was the most time consuming for both sexes, while cooking took 14 hours per week for working women (NISR, 2011)

The EICV3 found 479,000 people who said that they were self-employed in their main job and ran their own non-agricultural business, compared with the Establishment Census which found only 123,526 business and government establishments, employing 281,946 persons. This difference is primarily due to the EICV including smaller, informal 'firms', largely without employees or premises.

The ultimate goal of the study was to determine statistical linear model of which predictor variables accurately discriminate or separate women and men groups in wage non-farm employment sector in Rwanda.

1.2.1Justification of the study:

Discriminant Analysis Method was developed by Ronald Fisher in 1936, was the professor of statistics at the University College London, and is sometimes called Fisher Discriminant Analysis.

It was applied when Barbara A. Bardes, Phd wanted to know how USA Senate voting groups discriminated and changed over time and help them to know; how stable they were from year to year and how much they were influenced by other issues. This presents a need to apply the same research method in order to prove discrepancies between women and men of female regarding wage non-farm employment compared to their education level in Rwanda.

The NISR's databases of the Integrated Household Living Conditions Survey (EICV3) were useful. The study analysis required data mining which has been the process of selecting, exploring, and modeling large amounts of data to uncover new trends and patterns in massive databases. The 14,308 households were sampled units of the third EICV.

Discriminant Analysis works off of matrices used in Multivariate Analysis of Variance (MANOVA). Multiple quantitative attributes were used to build a linear model and to discriminate classification variables. The common objectives of DA : i) to investigate differences between groups ii) to discriminate groups effectively; iii) to identify important discriminating variables; iv) to perform hypothesis testing on the differences between the expected groupings; and v) to classify new observations into pre-existing groups.(Wiley & A, 1996)

Data was from distinct two groups. Group membership was known prior to initial analysis. Data did not have linear dependencies and was able to invert matrices. Population covariance was equal for each group. Each group was drawn from datasets of the 3rd EICV where the variables are multivariate normal.

1.3 Objective of the study:

The general objective of the study was to find out statistical linear model which predictor variables accurately discriminate or separate women and men groups in wage non-farm employment sector in Rwanda.

1.4 Specific objectives of the study:

1. To find out the statistical linear model of which predictors accurately separate male and female workers' groups.
2. To settle on predictor variables that are strongly related to groups discrepancies
3. To decide variables which are little related to groups distinctions

1.5 Research hypothesis:

1. There is no linear relationship between variables;
2. No predictor variables that are strongly related to group discrepancies
3. Some predictor variables are little related to groups distinctions

1.6 Significance of the study:

These analyses lead to proactive decision making and knowledge discovery in large databases by stressing data exploration to thoroughly study the structure of data and to check the validity of statistical linear models that fit. Linear Discriminant Analysis (LDA), a multivariate statistical technique was used to build a predictive model of group discrimination based on observed variables and to classify each observation into one of these two groups.

The findings of the study made known the significant separation of these two groups (women and men) against the nature of employment on the labor market to the Government policy makers, Private sector managers, Non-government organizations' managers, Civil society organizations' leaders, and Academic researchers. The expected research report took at the state of play from a gender perspective across all three issues, whether inequalities exist, how and why they have developed, and which obstacles need to be overcome to move towards greater equality. It offers policy advice to governments as to how they can create a more level playing field.

Gender policy, employment policy and strategies were reviewed and harmonized based on the expected findings of this study. The study shall allow statistical users and researchers to investigate further existing discrepancies between multiple sets of datasets across several variables simultaneously.

1.7 Scope of the study:

Statistical discrimination occurs when distinctions between demographic groups were made on the basis of real or imagined statistical distinctions between the groups (Dickinson, Oaxaca, & Dickinson, 2006). With purpose of getting two years' degree of masters in sciences in applied statistics, and the target population are female and male workers in wage non-farm employment sector in Rwanda, their type of wage non-farm employment and its predictors were very key variables of interests in this research.

1.8 Theoretical Framework:

Non-farming jobs in rural as well as urban areas currently relates to about 12.8 % of the Labor force that is approximately 433,596 people distributed in the various branches of economic activities. If we only consider the non-farming rural jobs they relate to only 1.8 % of the employed, that is 149,368 people in 2002. However, if this agricultural sector is well exploited, it will certainly create many jobs (Rwanda, 2007). Then the following is the theoretical framework of the study:

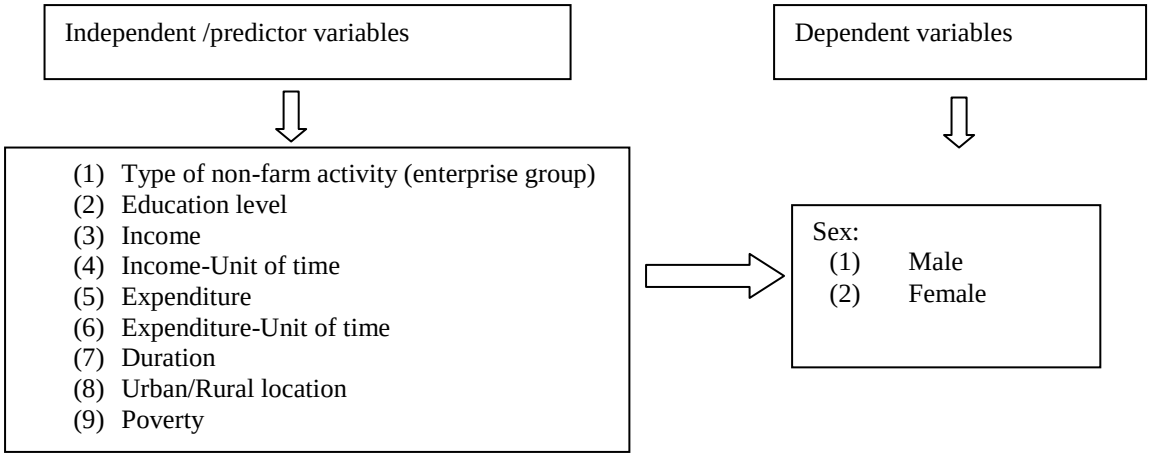

Figure 2: Theoretical framework

2. LITERATURE REVIEW

In 2011, Torben Iversen from the Havard University and Frances McCall Rosenbluth from the Yale University conducted a study on explaining Occupational Gender Inequality: Hours Regulation and Statistical Discrimination. Despite a large influx of women into mainly service sector jobs during the past four decades, women continue to be under-represented in the labor market, and they earn less on average than men. These gender differences are almost certainly linked to greater de facto responsibilities of women in child rearing and household work, but there are major and intriguing differences across rich democracies (Torben & Frances, 2011).

In 1993-94 under the supervision of Professor Michael Lipton, Dr. Diana Hunt and the late Dr. Pramit Chaudhuri from the University of Sussex, Prasada Rao Mecharla explained the factors which affect rural non-farm employment in two villages; using primary data from the Indian state of Andhra Pradesh, analyzed the reasons for the variations between an agriculturally-developed village and one which is less developed. The survey, conducted during 1993-94 covered a total of 465 households in Veeravalli and 354 households in Anandapuram villages of two districts of Andhra Pradesh. In this region of India there is of course the multivariate effect of farm size; because there are higher levels of production following from ownership of more land, the production-consumption linkages mean that the likelihood of non-farm employment is increased through another channel. Wealthier villagers invest more in their children's education, which increases the likelihood of them taking non-farm employment; they also consume more, meaning more work for others. In the pooled data set one more year in education increases the chances of getting modern non-farm employment by 5 percentage points. It seems some of the results have economic and non-economic barriers. Education is significantly positively related to non-farm employment in each village, and on the pooled data. The marginal effect of education on the probability of HH having non-farm employment is greater in the more developed village compared to the less developed village. In this case, the effect of literacy in raising demand and/or supply for 'modern' RNFE appears to outweigh the effect in reducing them for traditional RNFE (Prasada, 2002).

The Oaxaca-Blinder decomposition technique was developed independently by these two economists in 1973, and has since been elaborated on by Cotton (1988) and Neumark (1988) among others. Its virtue lies in allowing for the possibility that discrimination might be reflected not just in a fixed differential between the wages of, for example, men versus women, but also by differences in the rewards associated with increases in men's versus women's human capital. The wisdom of this observation has recently been affirmed in an analysis of hiring (as opposed to wage) discrimination against African-Americans, which found that educational credentials are heavily discounted for blacks, and that this explains a portion of their lower interview call-back rate by employers who, in a randomized experiment (Thomas, Paul, & Alberto, 2008).

In Nebraska from 1977 to 1985 Jane C. Ollenburger, Sheryl J. Grana, and Helen A. Moore analyzed the paid labor force participation rates and continuity patterns of rural farm, rural nonfarm, and urban women. Specifically, we trace the labor force participation of a panel of approximately 800 women in Nebraska from 1977 to 1985. Their paper has a twofold purpose. First, they examine changes in the work status of the cohort of Nebraska women during the farm crisis years. Second, they identify individual factors influencing labor force participation and continuity, contrasting all three

residential groups of women. A log-linear model isolates differences in participation rates for rural and urban women as well as for rural farm and rural nonfarm women during the 1977, 1981, and 1985 panel years. A discriminant analysis then ascertains the nonlinear relationships in women's work histories for the same time period. A comparison of continuous, discontinuous, and non-participatory labor force patterns illustrates that rural women and farm women in particular, entered the wage labor force in disproportionate numbers during the farm crisis years. The farm crisis provides a framework for discussing accelerated participation rates and changes in the effects of individual human capital characteristics. Increases in participation rates are most evident among married farm women with discontinuous part time work histories. Over the three data points, the effect of preschool children on labor work force participation was attenuated for farm women and higher education levels became less salient in predicting labor force participation rates for both rural and urban women (Jane, Sheryl, & Helen, 1989).

In 2012, Milena PACCHIOTTI from the Institut d'études politiques de Paris (Sciences Po) conducted analysis of the multifaceted dimensions - economic, legal, statistical - of labour market gender equality, with particular focus on the situation in Tanzania. This research reveals that seven out of ten employed females work on their own farm compared to six males out of ten, confirming that agriculture is the main sector of activity of women. Men are more represented as paid employees and self-employed in non-agricultural activities, which are also more profitable sectors. The percentage of male and female as unpaid family helpers in agriculture is quite balanced. Once again, the largest disparity is founded in non-agricultural activities, where female unpaid family helpers are 5 times more than their male counterparts. This evidence allows supposing different interpretations of gender roles in rural and urban areas that attributes distinct workload to women in households. Women in urban areas who are involved in non-agricultural activities, are more likely than men to work as unpaid family helpers. On the other hand, women in rural areas work more than men as self-employed in their own farm (Milena & Pablo, 2012).

In 2014, Gloria Guangye He and Xiaogang Wu This published an article that examines the changing trends in gender earnings inequality—an important aspect of social stratification—in the context of macro-level social and economic changes, with special attention paid to the differential impacts in shaping the socioeconomic relationship between men and women in urban China. They used data from the 2005 population mini-census and prefecture-level statistics. They pay special attention to the different impacts of marketization and economic development on gender earnings inequality. Cross-sectorial analyses show that the gender earnings gap is smallest in government and public institutions and increases for more marketed sectors. At the prefectural level, they match the mini-census data with prefecture-level statistics and differentiate the effect of economic development from that of marketization. Multi-level analyses show that marketization and economic development affect gender inequality in different ways: the market force has exacerbated gender earnings inequality, whereas economic development has reduced it. Overall, marketization appears to be the main driver of the increasing gender earnings inequality in urban China. Findings shed new light on the changing gender inequality and the effective policies to promote gender equality in urban China's labor markets. The gender earnings inequality is mainly attributable to women's relatively lower education and occupational gender segregation (Gloria & Xiaogang, 2014).

Although the importance of reducing gender discrimination for social and economic development is now widely recognized, most studies focus on the formal sector. Representative data from rural India demonstrate that informal markets, which are of great relevance for the poor, are characterized by high gender-wage differentials that are unrelated to productivity. While increasing wage levels, economic development fails to reduce such gender discrimination. Large differences in estimated marginal value products between family and hired labor suggest that improving access by the poor to productive assets that could facilitate self-employment may have significant benefits (Klaus, Songqing, & Hari, 2006).

3. METHODOLOGY

I was interested in the relationship between a group of independent variables and one categorical variable. I would like to know how many dimensions would be needed to express this relationship. Using this relationship, I predicted a classification based on the independent variables or assess how well the independent variables separate the categories (two classes: women and men in wage non-farm employment.

Data source for analysis was the 3rd EICV published by the National Institute of Statistics of Rwanda (NISR). Knowing that the 3rd EICV revealed that working age people are 16 years old and plus because many, especially in the 14–19 age category, have been still students.

The 2010/11 Integrated Household Living Conditions Survey or EICV3 (Enquête Intégral sur les Conditions de Vie des Ménages) was the third in the series of surveys which started in 2000/01 and was designed to monitor poverty and living conditions in Rwanda. The survey methodology has changed little over its 10 years, making it ideal for monitoring changes in the country. In 2010/11, for the first time the achieved sample size of 14,308 households in 1,230 sampled villages in the EICV3 was sufficient to provide estimates which are reliable at the level of the district and great number of observations were collected, processed and published as raw data by the NISR in 2012. Our concern will be the analysis of earnings and occupations of both female and male group workers in Rwanda. The following steps were followed:

3.1 The organization of data:

Issues of costs and time limitations push me to use 3rd EICV raw data, a huge package of numerical observations were collected and ready for deep analysis. The sample of 14,308 individuals is large enough to be multivariate normal distributed. Then the test of normality was not needed.

In small samples, values greater or lesser than 1.96 are sufficient to establish normality of the data. However, in large samples (200 or more) with small standard errors, this criterion should be changed to ± 2.58 and in very large samples no criterion should be applied (that is, significance tests of skewness and kurtosis should not be used) (Field, 2009)

Throughout this text, I was concerned with analyzing measurements made on independent variables. These measurements shall be frequently arranged and displayed in various ways. Thus, SPSS tabular arrangements will be important aids in data organization, summary numbers, which will quantitatively portray certain features of the data, will be also necessary to the data description. Then, selects a number $p \geq 1$ of variables or characters to record. The values of these variables are all recorded for each distinct individual or respondents of Integrated Households Living Conditions Survey (EICV 2010-2011) conducted by National Institute of Rwanda (NISR). I will use the notation x_{jk} to indicate the particular value of the k^{th} variable that is observed on the j^{th} item, or trial. That is, x_{jk} = measurement of the k^{th} variable on the j^{th} item

3.2 Linear Discriminant Modelling:

Linear Discriminant Analysis (LDA) was a classification method originally developed in 1936 by R.A. Fisher. It was simple, mathematically robust and often produces models whose accuracy was as good as more complex methods. It was based upon the concept of searching for a linear combination of variables (predictors) that best separates two groups (targets). To capture the notion of separability, Fisher defined the following score function:

$$z = \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_d x_d$$

For my study

$$\hat{z} = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7 + \beta_8 x_8 + \beta_9 x_9 + \varepsilon$$

Assumptions: $\mathrm{E}(\varepsilon) = 0$, $Var(\varepsilon) = \sigma^2$, ε are independent and identically distributed, $\varepsilon \sim N(0, \sigma^2)$

Where

$\hat{z}$:The target two groups (men and women), x_1: Occupation group, x_2: Education level, x_3: Income, x_4: Income-Unit of time, x_5: Expenditure, x_6: Expenditure-Unit of time, x_7: Duration, x_8: Urban/Rural location, x_9: Poverty. On behalf of the outcome (z) data were categorical/binary and for predictors (x_i) data were numerical. In order to estimate linear coefficient, score function is given as

$$s(\beta) = \frac{\beta^T \mu_1 - \beta^T \mu_2}{\beta^T C \beta}, \quad s(\beta) = \frac{\overline{z_1} - \overline{z_2}}{\text{var } z_w}$$

Where

$s(\beta)$: Score function, μ_1: The mean of the subset for men, μ_2: The mean of the subset for women, β^T : Transpose vector of coefficients, $\overline{z_1}$: The mean of the group of men, $\overline{z_2}$: The mean of the group of women, C: The covariance matrix. Given the score function, the problem is to estimate the linear coefficients that maximize the score which can be solved by the following equations:

- Model coefficients: $\beta = C^{-1}(\mu_1 - \mu_2)$
- Poole covariance matrix: C= $\frac{1}{n_1 + n_2}(n_1 C_1 + n_2 C_2)$

Where β : Linear model coefficients, C_1, C_2 : Covariance matrices, μ_1, μ_2 : Mean vectors.

One way of assessing the effectiveness of the discrimination is to calculate the mahalanobis distance between two groups. A distance greater than 3 means that in two averages differ by more than 3 standard deviations. It means that the overlap (probability of misclassification) is quite small.

$$\Delta^2 = \beta^T(\mu_1 - \mu_2)$$

Δ: Mahalanobis distance between group of men and group of women

Finally, when a new point comes in, it may be classified by projecting it into the maximally separating direction and classifying it as C_1 if:

$$\beta^T\left(x - \left(\frac{\mu_1 + \mu_2}{2}\right)\right) > \log\frac{p(C_1)}{p(C_2)}$$

Where β^T : Coefficients vector, $p(C_1)$: Class probability of the men group, x : New data vector $p(C_2)$: Class probability of the women group, μ_1, μ_2 : Mean vector.

3.3 Proposed statistical tools and tests:

The principal tool to be used in the study was the SPSS. It was powerful open source software which can help to carry out various statistical analyses. It will be used to undertake the various statistical tests and analysis. The tests include the Univariate ANOVA test, the Box's M test, the Wilks' Lambda test and Canonical Discriminant Analysis.

3.4 Data analysis and procedure:

The data to be used in the research will be obtained from National Institute of Statistics of Rwanda (NISR). Data are available and sourced from the 3rd EICV that was conducted by the National Institute of Statistics of Rwanda (NISR) in 2010-2011.

It is intended that raw data will be in SPSS. Once the data is obtained, a data cleaning exercise will be embarked on. In case of missing data points, interpolation will be applied to fill these gaps. Then the analysis will be completed.

4. RESULTS AND DISCUSSIONS

4.1. Data description, source, period and nature of data:

This chapter has examined the best predictors in the model strongly and or weakly separate men and women in the non-farm employment sector in Rwanda. These predictors are type of non-farm employment, education, poverty, income, and income unit of time, expenditure, expenditure unit of time, duration and urban/rural location. For dependent variable (men and women groups) data are quantitative and categorical binary, and for predictors data are numerical.

The data used for the analysis was taken from the 2010/11 Integrated Household Living Conditions Survey or EICV3 (Enquête Intégrale sur les Conditions de Vie des Ménages) that was the third in the series of surveys which started in 2000/01 and was designed to monitor poverty and living conditions in Rwanda. The survey methodology has changed little over its 10 years, making it ideal for monitoring changes in the country. In 2010/11, for the first time the achieved sample size of 14,308 households in the EICV3 was sufficient to provide estimates which are reliable at the level of the district. I took this period of analysis as quick alternative as raw data of EICV 4 have not yet been available until September 2016. By EICV 3 questionnaire were administered on 61,405 individuals from 14,308 households sampled, in 1,230 selected villages of all 30 districts of Rwanda. The research's sample has been limited to the age between 18 and 65 years old by which women and men who are engaged in the wage nonfarm employment sector. This implied that 7,353 individuals belonged to the actual sample size with 3,772 (51.3%) women and 3,581 (48.7%) men. Among 55,613 respondents who were aged between 18 and 65, only 7,353 are those who were engaged in the wage nonfarm employment sector with 49% of men against 51% of women.

4.2 Demographic characteristics:

In this part, the demographic and background information of the NFE workers are presented and analyzed in order to show the distribution of the respondents by their sex and age.

Table 2: Non-Farm Job categories between women and men

Occupation – Grouped	Male	%	Female	%	Total
Professionals	116	89%	14	11%	130
Senior Officials and Managers	5	100%	0	0%	5
Office Clerks	13	68%	6	32%	19
Commercial and Sales	2,067	45%	2,557	55%	4,624
Skilled Service Sector	127	56%	101	44%	228
Agricultural & Fishery Workers	79	89%	10	11%	89
Semi-Skilled Operatives	870	46%	1,039	54%	1,909
Drivers and Machine Operators	294	97%	8	3%	302
Unskilled Labourers	42	89%	5	11%	47
Total	3,613	49%	3,740	51%	7,353

Source: Primary data analysis for linear discriminant modelling of wage non-farm employment between women and men in Rwanda

Professionals are 130 out of 7,353 with 89% men and 11% women, drivers and machine operators are 302 out of 7,353 by 97% men and 3% women, skilled service providers are 228 out of 7,353 with 56% men and 44% women, semi-skilled operators are 1,909 out of 7,353 with 46% men and 54% women, and commercial and sales operatives are 4,624 men and women with 45% and 55% respectively. The first large numbers of people in the NFE are commercial and sales operatives with 63% (4,624) of all. The second large numbers of these people are semi-skilled operatives by 26% (1,909) of all.

Table 3: Sex disaggregated of women and men by urban and rural province

	Sex		Location		Province				
	Nber	**%**	**Urban**	**Rural**	**Kigali City**	**South**	**West**	**North**	**East**
Male	3,581	48.7%	51%	48%	53%	48%	47%	48%	50%
Female	3,772	51.3%	49%	52%	47%	52%	53%	52%	50%
Total	7,353	100%	100%	100%	100%	100%	100%	100%	100%

Source: Primary data analysis for linear discriminant modelling of wage non-farm employment between women and men in Rwanda

People were asked to mention local residence area by then urban area covered 51% of men and 49% of women while in the rural area 48% and 52% are men and women respectively. It was evident that majority of men are working in the urban areas whereas majority of women are in the rural areas.

Table 4: Age intervals for men and women

Age Interval	Male	%	Female	%	Total	%
[18-22]	554	49%	576	51%	1,130	15%
[23-27]	693	50%	693	50%	1,385	19%
[28-32]	649	51%	624	49%	1,273	17%
[33-37]	423	48%	458	52%	881	12%
[38-42]	347	47%	391	53%	737	10%

[43-47]	291	49%	303	51%	594	8%
[48-52]	262	46%	308	54%	570	8%
[53-57]	187	46%	219	54%	406	6%
[58-65]	170	45%	208	55%	377	5%
Total					7,353	100%

Source: Primary data analysis for linear discriminant modeling of wage non-farm employment between women and men in Rwanda

The above table shows that majority of respondents are between 18 and 32 years old. The age intervals of [18-22] and [23-27] men are 49% of 1,130 and 50% of 1,385 respectively. It is obvious that youth population is growing fast in the wage nonfarm employment sector in Rwanda.

According to the fourth Rwanda General Population and Housing Census (RGPHC 4), the active population, which consists of employed and unemployed persons, represented 74% of the population aged 16 and above. Females were predominant in the active population and also constituted the majority of the working-age population. Compared to this study in the NFE, 63% are the young generation of less than 37 ages who are enormously engaged in the NFE.

Table 5: Marital status for women and men

Marital status	**Sex**				**Total**	**%**
	Male	**%**	**Female**	**%**		
Married monogamously	1908	54%	1625	46%	3533	48%
Married polygamously	101	55%	83	45%	184	2%
Living together	507	54%	432	46%	938	13%
Divorced	2	21%	6	79%	8	0.1%
Separated	43	15%	243	85%	286	4%
Single	970	52%	895	48%	1865	25%
Widow or widower	32	6%	507	94%	540	7%
Total					7353	100%

Source: Primary data analysis for linear discriminant modelling of wage non-farm employment between women and men in Rwanda

The above table proves that the number of monogamously married people is near to the half of all respondents with 48%. Single status occupies 25% of all workers in the NFE. Generally, the number of women is higher than the number of men but especially in the divorced and separated status with which women stand at 79% and 85% respectively. As conclusion, monogamously married people are the majority of all workers in the NFE in Rwanda.

Table 6: Education attainment by sex

	Observation	Male	%	Female	%	Total	%
Have you ever been to school?	Yes	3,000	51%	2,882	49%	5,882	80%
	No	559	38%	912	62%	1,471	20%
	Total	3,559	49%	3,794	51%	7,353	100%

Source: Primary data analysis for linear discriminant modelling of wage non-farm employment between women and men in Rwanda

People were asked if ever he or she had been to school, yes response accommodates around 5,882 (80%) out of 7,353, 51% of them are men and 49% of them are women. However, no response was given by 1,471 (20%) out of 7,353 with 38% and 62% of men and women respectively.

Table 7: School completion levels for men and women

Education completed level	Men	%	Female	%	Total	%
Primary completed	2373	52%	2162	48%	4534	77,09%
Secondary common	160	58%	115	42%	274	4,66%
Post primary certificate	195	54%	168	46%	364	6,18%
Diploma A3, D5, D4	33	61%	21	39%	53	0,91%
Humanities Diploma	250	54%	215	46%	465	7,91%
Bachelors	30	65%	16	35%	45	0,77%
Professional license	73	65%	39	35%	112	1,91%
Engineer	6	85%	1	15%	7	0,13%
Masters	14	67%	7	33%	21	0,36%
Doctorate	5	93%	0	7%	5	0,09%
Total	3138	53%	2744	47%	5882	100%

Source: Primary data analysis for linear discriminant modelling of wage non-farm employment between women and men in Rwanda

In the NFE, majority of people engaged in this sector have completed primary level school with 77% of all participants in the sector. The level of Diploma is at 1%, Bachelor with 0.7%, Masters, Engineering, and Doctorate with 0%. It is evident that non-high education level in the sector what is really expected? It is almost nothing rather than routine employment without innovations.

Table 8: Poverty status for men and women

Poverty status	Sex				Total	%
	Male	%	Female	%		
Extremely poor	744	44%	947	56%	1,691	23%
Poor	726	47%	818	53%	1,544	21%
Non-poor	2,100	51%	2,018	49%	4,118	56%
Total	3,570	49%	3,783	51%	7,353	100%

Source: Primary data analysis for linear discriminant modelling of wage non-farm employment between women and men in Rwanda

It has been that 5,910 (80%) out of 7,353 workers in the NFE had gone to school, 23% of them are extremely poor, 21% who are poor and 56% non-poor workers. By all these categories women are major group with 51% against 49% of men. It is mentioned that many of them are living in the rural areas.

4.3 Results of Linear Discriminant Analysis (LDA):

In the Non-Farm Employment sector, LDA was used to provide the maximum separability between two groups (men and women groups) which has been a categorical variable in linear relationship with a group of predictor variables. By following objective sequence, research assumptions were tested to different forms of LDA.

Objective 01: To determine the statistical linear model of which predictors accurately separate male and female workers' groups.

Ho: There is no linear relationship between variables

Ha: There is linear relationship between variables

4.4 ANOVA Test of Non linear relationship between variables:

We applied hypothesis test to determine whether there is a significant linear relationship between independent variables x and a dependent variable *y*. rule: If the sample findings are unlikely, given the null hypothesis, the researcher rejects the null hypothesis. Typically, this involves comparing the P-value to the significance level (0.05), and rejecting the null hypothesis when the P-value is less than the significance level (refer to the page 16).

Then, We have

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7 + \beta_8 x_8 + \beta_9 x_9 + \varepsilon$$

Where y : Sex (men and women), α : Constant value, x_1 : Occupation Group (OG), x_2 : Education Level (Edu), x_3 : Income (Inc), x_4 : Income-Unit of Time (IUT), x_5 : Expenditure (Exp), x_6 : Expenditure-Unit of Time (EUT), x_7 : Duration (Dur), x_8 : Urban/Rural Location (URL), x_9 : Poverty (Pov), ε : Error value/residuals.

Table 9: Results of the ANOVA test of no linear relationship between variables

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	82.094	10	8.209	34.555	.000[b]
Residual	1516.904	6385	.238		
Total	1598.997	6395			

Source: Primary data analysis for linear discriminant modelling of wage non-farm employment between women and men in Rwanda

Table 10: Distribution of coefficients for predictor variables

Independent variables	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	1.864	.052		36.014	.000
Education level	.165	.018	.115	9.203	.000
Occupation - Grouped	-.031	.004	-.096	-7.571	.000
Industry Group	2.999	.000	.079	6.351	.000
Expenditure on labour	5.275	.000	.061	3.449	.001
Income of business	-3.211	.000	-.117	-6.513	.000
Income Unit of time	-.020	.009	-.035	-2.154	.031
Expenditure Unit of time	-.009	.008	-.019	-1.174	.240
Duration business functioning-years	.001	.001	.008	.616	.538
Poverty status	-.074	.009	-.106	-8.452	.000
Urban/rural	-.077	.016	-.061	-4.892	.000

Source: Primary data analysis for linear discriminant modelling of wage non-farm employment between women and men in Rwanda

The P-value is less than 0.05 for a series of independent variables (Edu, OG, Inc, IUT, Exp, Pov, and URL). This means that there is strong evidence to reject null hypothesis, and accept that variables are linearly combined. However, duration and expenditure unit of time have p-value above 0.05 which allows to not reject the null hypothesis. But generally the p-value of regression is less than 0.05 to prove that the function is linear statistically significant.

Then, we have linear equation:

$y =$ 1.864 + (-0.31) *OG + 0.165*Edu + (-3.211) *Inc + (-0.020) *IUT + 5.275*Exp + (-0.009) *EUT + 0.001*Dur + (-0.077) *URL + (-0.074) *Pov + 1516.904

As we add one unit of increase to each independent variable, the dependent variable increases or reduces by times of coefficient value.

4.5 Canonical discriminant test of correlation:

An Eigen value indicates the proportion of variance explained. (Between-groups sums of squares divided by within-groups sums of squares). A large Eigen value is associated with a strong function. When the projection vector corresponding to the smallest Eigen value, using this vector leads to bad separability between the two groups. However, the projection vector corresponding to the highest Eigen value, using this vector leads to good separability between the two groups. And the canonical relation is a correlation between the discriminant scores and the levels of the dependent variable. A high correlation indicates a function that discriminates well variables.

Different levels of analysis were applied as follow:

Ho: Correlation exists

Ha: Correlation does not exist

Table 11: Results about canonical correlation – Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	.016[a]	100.0	100.0	.127

Source: Primary data analysis for linear discriminant modeling of wage non-farm employment between women and men in Rwanda

The canonical relation is described by a correlation between the discriminant scores and the levels of the dependent variable. This correlation value has been squared (0.127^2 = 0.016129). It is equal to eigenvalue of 0.16 which is the higher value that best provide a strong function and a good separability between group of men and group of women.

Table 12: Test of variance within group difference – Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	Df	Sig.
1	.984	906.183	10	.000

Source: Primary data analysis for linear discriminant modelling of wage non-farm employment between women and men in Rwanda

Ho: Variance within groups is not significant

Ha: Variance within group is significant

The associated significance value is 0.000 which is less than 0.05. This indicates that variance within group difference is not significant. This implies that further researches are needed to identify factors that can sufficiently influence variation within groups (men and women).

Table 13: Standardized canonical discrimination and structured matrix

Standardized Canonical Discriminant Function Coefficients

	Function 1
Occupation group	-.253
Expenditure	.178
Exp. Unit of time	-.087
Income	-.331
Income unit of time	-.032
Industry group	.178
Duration	.044
Urban/rural	-.012
Poverty status	-.350
Ever been to school	.808

Structure Matrix

	Function 1
Ever been to school	.853
Poverty status	-.449
Income	-.213
Occupation group	-.194
Industry group	.154
Exp. Unit of time	-.146
Income unit of time	-.131
Urban/rural	.115
Expenditure	-.063

Source: Primary data analysis for linear discriminant modelling of wage non-farm employment between women and men in Rwanda

The above canonical discriminated function coefficients indicate the standardized scores concerning the independent variables. It is the list of coefficients of the standardized discriminant equation. Each subject's discriminant score has been computed by entering variable values (raw data x column in the SPSS) for each of the variables in the equation. Rule: if the variable value is greater than 0.3, the variable intends to maximally separate groups and if it is less than 0.3, the variable intends to minimally separate groups. Thus the structured matrix gives coefficients' scores to Linear Discriminant Model as follow:

Z = 0.853*Edu + (-0.449) *Pov + (-0.213) *Inc + (-0.194) *OG + 0.154*Ind + (-.146) *EUT + (-.131) *IUT + 0.115*URL + (-0.63) *Exp + 0.61*Dur

It is clear that one unit of increase in the poverty, income, occupation group, income unit of time, expenditure unit of time and expenditure explains little to discrimination between men and women in the NFE sector in Rwanda while one unit of increase in education, industry activities, urban/rural location and duration in business explains better and maximally discrimination between women and men in the NFE sector in Rwanda.

4.6 Test of equality of group means:

A test for the equality of the group covariance matrices. For sufficiently large samples, a non-significant p value means there might be insufficient evidence that the matrices differ in case that wilks' lambda equals to 1 but when it becomes less than 1, group means differ for two groups.

Ho: all means are equal, **Ha:** all means are not equal

Table 14: Tests of equality of group means

Independent variables	Wilks' Lambda	F	df1	df2	Sig.
Occupation group	.999	34.427	1	55611	.000
Expenditure	1.000	3.591	1	55611	.058
Exp. Unit of time	1.000	19.464	1	55611	.000
Income	.999	41.548	1	55611	.000
Income unit of time	1.000	15.740	1	55611	.000
Industry group	1.000	21.533	1	55611	.000
Duration	1.000	3.421	1	55611	.064
Urban/rural	1.000	12.052	1	55611	.001
Poverty status	.997	183.908	1	55611	.000
Ever been to school	.988	664.247	1	55611	.000

Source: Primary data analysis for linear discriminant modelling of wage non-farm employment between women and men in Rwanda

The above table shows that Wilks' lambda is equal to 1 for six predictor variables (expenditure, expenditure unit of time, income unit of time, industry group, duration, and urban/rural location). This means that group means do not differ for both men and women consequently all variance within group is explained by factors other than difference between those means. Therefore, means are not sufficiently enough to explain the variation within group of men and group of women in NFE sector in Rwanda. For the rest of variable (occupation group, income, poverty status and education), the lambda is less than 1 meaning that group means differ for both men and women.

Objective 02: To determine predictor variables that are strongly related to groups discrepancies

Ho: There are no predictor variables that were strongly related to group discrepancies

Ha: There are predictor variables that were strongly related to group discrepancies

4.7 Classification Statistics:

By classification, we discuss about prior probability for groups, classification function coefficients, and classification results for group membership.

Table 15: Prior probabilities for groups

Sex	Prior	Unweighted cases	Weighted cases
Male	.487	3578	3578.000
Female	.513	3775	3775.000
Total	1.000	7353	7353.000

Source: Primary data analysis for linear discriminant modelling of wage non-farm employment between women and men in Rwanda

According to group distinction men are 48.7% while women are 51.3% in the wage nonfarm employment. This is almost the same percentage of the results of the fourth Rwanda General Population and Housing Census (RGPHC 4) by which 51% stands for women and 49% of men.

Table 16: Classification function coefficients

Independent variables	Male	Female
Occupation group	13.707	13.593
Expenditure	1.830E-7	4.148E-7
Exp. Unit of time	11.077	11.014
Income	2.076E-7	7.034E-8
Income unit of time	13.608	13.583
Industry group	-.002	-.002
Duration	-.014	-.010
Urban/rural	17.620	17.610
Poverty status	5.562	5.453
Ever been to school	7.369	7.893
(Constant)	-93.788	-93.317

Source: Primary data analysis for linear discriminant modelling of wage non-farm employment between women and men in Rwanda

As given by the above table, there are two linear discriminant functions; one is male function and another is female function. The function coefficients values which are above 0.3 assign strong relationship to its predictors on dependant variable while those with values below 0.3 correspond to weak relationship with dependant variable. For this case, only two weak predictors are industry group and duration. The remaining predictors are considered to be strongly related to men group and women group. The constant value for men group's linear function is -93.788 and -93.317 for women's linear function.

Let $\hat{z}_M$ denote the estimated responses for males and $\hat{z}_F$ be the estimated responses for females

$\hat{Z}_M$ = -93.788 + 13.707*OG + 1.830*Exp + 11.077*EUT + 2.076*Inc + 13.608*IUT + (-.002) *Ind + (-.014) *Dur+ 17.620*URL + 5.562*Pov + 7.369*Edu

$\hat{Z}_F$ = -93.317 + 13.593*OG + 4.148*Exp + 11.014*EUT + 7.034*Inc + 13.583*IUT + (-.002) *Ind + (-.010) *Dur + 17.610*URL+ 5.453*Pov + 7.893*Edu

The case is the same for both men and women. Values of coefficients or weights of variables like occupation group, expenditure, expenditure unit of time, income, and income unit of time, urban/rural habitation, poverty status and education, are valued greater than 0.3. They explain that discrimination between men and women in NFE sector is strongly or maximally related to those variables. In addition, if we increase one unit in the above independent variables, discrimination weight shall increase by times values of coefficients for more and more between men and women. Then the null hypothesis is failed to be true.

Objective 03: To discard variables which are little related to groups distinctions

Ho: There are no predictor variables that are little related to groups distinctions

Ha: There are predictor variables that are little related to groups distinctions

If we linked this objective to the above statement, we get from the model two independent variables (duration in the business and industry group of jobs) that are likely to be weak or little related to the discrimination between women and men in the NFE sector in Rwanda. these variables have coefficients values that are below 0.3. It is evident that duration in the job or business and job categories are lowly influencing factors of men and women differences and the null hypothesis is rejected.

5. CONCLUSION, ACKNOWLEDGEMENT AND RECOMMENDATION

Conclusion and Acknowledgement:

The study was through deep analysis of secondary raw data collected in the EICV 3 between 2009-2011. The NFE was the sector of interest among other sectors in the economic activities for its relevance in the economic growth of Rwanda. It is the one that provide more inputs in the agriculture and services sectors. The study shows that 7,353 individuals are engaged in the NFE with 51% of men and 49% of women. Though men are predominant in the sector but the difference is yet low to be significant. The level of education among the NFE workers is still very weak. The study shows that among 80% of the educated active population in the NFE, 77% of them completed only primary level of education.The study reveals that inequality is still acceptable in the NFE sector where occupation group, expenditure, expenditure unit of time, income, and income unit of time, urban/rural habitation, poverty status and education are the strong determinants of the maximum discrimination between men and women in the NFE in Rwanda.

I would like to express my sincere gratitude to my supervisors, Joseph K. MUNGATU, Phd and Marcel NDENGO, Phd for their constant support, guidance and constructive ideas all over my research project. Special thanks go to the Administration Staff members of Jomo Kenyatta University of Agriculture and Technology Campus of Kigali for their support throughout this work processes. I would like to thank my fellow master's students in the Department of Applied Statistics. Their cooperative spirit and contribution during the whole period of my study is appreciated. Last but not least, I would like to express my thanks to my wife, brothers and sisters for their love and encouragement during the whole period of my study.

Recommendations:

Against the above statements, I recommend the following:

1. The Government of Rwanda should encourage women in NFE to increase their education level and put in place strategies to support vocational trainings in favor of NFE sector;
2. The Government of Rwanda in partnership with stakeholders should increase the number of women in the NFE sector;
3. Academic researches should be supported to search and find out reasons behind inequalities in terms of expenditure and linked time factors that separate women and men;
4. Men and women must be aware of job opportunities in the NFE sector and well manage the income gained from them;
5. The Private sector under support of the Government of Rwanda should create more jobs in the urban as in the rural areas for equal opportunities so as to minimize difference between men and women;
6. Further researches should be carried out in the areas of non-farm employment sector in Rwanda and in the East African Region.

REFERENCES

[1] Abdi H. (2007). Discriminant correspondence analysis. N.J. Salkind. Encyclopedia of Measurement and Statistic

[2] Aliyari, G. and Rudzicz, F. (2015). Fast incremental LDA feature extraction.

[3] Ahdesmäki, M. and Strimmer, K. (2010). Feature selection in omics prediction problems using cat scores and false nondiscovery rate control. Annals of Applied Statistics.

[4] Bardes, B. (2005). Public Opinion: Measuring the American Mind. USA.

[5] AfDB. (2014). Analysis of Gender and Youth Employment in Rwanda.

[6] Carletto, G., and Katia, C. (2007). Rural income generating activities in developing countries. Electronic Journal of Agricultural and Development Economics, 4.

[7] Demeke, M., Guta, F., and Ferede, T. (2006). Towards a more employment-intensive and pro-poor economic growth in Ethiopia: Issues and Policies, Issues in Employment and Poverty.

[8] Dickinson, D. L., Oaxaca, R. L., and Dickinson, D. L. (2006). Statistical Discrimination in Labor Markets : An Experimental Analysis, (2305).

[9] Field, A. (2009). Discovering statistics using SPSS.

[10] Gloria, G. H., and Xiaogang, W. (2014). Gender Earnings Inequality in Reform-Era Urban China. Hong Kong University of Science and Technology, 14-830.

[11] GoR, and One UN. (2013). Youth and Women Employment.

[12] HUNG, T. P. (2006). Trade liberalization and rural nonfarm employment in Vietnam. Vietnam

[13] Jane, C. O., Sheryl, J. G., and Helen, A. M. (1989). Labor Force Participation of Rural Farm, Rural Nonfarm, and Urban Women. University of Minnesota.

[14] Jütting, J., Luci, A., and Morrison, C. (2010). Why Do So Many Women End Up in "Bad Jobs" ? A Crosscountry Assessment for Developing Countries. OECD Development Centre Working Paper, 287.

[15] Kati, S. (2008). Time allocation, gender and norms: evidence from post-genocide Rwanda.

[16] Klaus, D., Songqing, J., and Hari, K. N. (2006). Gender discrimination and returns to self-employment Evidence from rural India. World Bank.

[17] Malunda, D. (2012a). Employment intensity of non-agricultural growth in Rwanda: Analyzing the links between Growth, Employment, and Productivity. IPAR, Rwanda.

[18] Malunda, D. (2012b). Report for the African Centre for Economic Transformation: Case of Rwanda. IPAR.

[19] Mary, S., Yussuf, M., and Emily, W. (2011). Labor market and economic activity trends in Rwanda.

[20] Milena, P., and Pablo, S. R. (2012). Gender (in)equality in the Tanzanian labour market: showing the gap between the legal framework and the evidence provided by labour statistics.

[21] Ministry of Youth. (2010). Rwanda Youth Statistical Indicators. Republic of Rwanda.

[22] MoH. (2015). Adolescent sexual reproductive health and rights policy. Rwanda: Ministry of Health.

[23] Murenzi, I., Pamela, A., and Musana, S. (2012). The Role of Non-farm Household Enterprises in Poverty Reduction, Employment Creation and Economic Growth in Rwanda. Rwanda Journal, 26 series B.

[24] NISR. (2007). EICV poverty analysis for Rwanda's Economic Development and Poverty Reduction Strategy.

[25] NISR. (2011). Integrated Economic Living Conditions Household Survey (EICV 3).

[26] NISR. (2012a). Integrated Economic Household Living Conditions (EICV 3): Musanze District Profile.

[27] NISR. (2012b). The evolution of poverty in Rwanda from 2000 to 2011: results from the household surveys (EICV).

[28] NISR, R. (2013). Thematic Report on Gender.

[29] OECD. (2012). Gender Equality in Education, Employment and Entrepreneurship. Paris.

[30] Prasada, R. M. (2002). The determinants of rural non-farm employment in two villages of Andhra Pradesh (India). University of Sussex, 02.

[31] Roger, N., Dan, G., and John, P. (2014). Staff report for the 2014 article IV consultation and second review under the policy support instrument—debt sustainability analysis.

[32] Rosabeth Moss, K. (2008). Men and Women of the Corporation. PublicAffairs.

[33] Rwanda, M. (2007). National Employment Policy.

[34] Rwanda, N. (2010). Rwanda Demographic and Health Survey.

[35] Rwanda, N. (2012). General Population and Housing Census. National Institute of Statistics of Rwanda.

[36] Thomas, H., Paul, W., and Alberto, Z. (2008). Wage Inequality in International Perspective: Effects of Location, Sector, and Gender. UN FAO.

[37] Torben, I., and Frances, M. R. (2011). Explaining Occupational Gender Inequality: Hours Regulation and Statistical Discrimination. Havard University & Yale University.

[38] UN. (2015). Economic Diversification, Non-Farm Activities and Rural Transformation: The Least Developed Countries Report, UN Conference on Trade and Development.

[39] USAID/Rwanda. (2015). Gender Analysis for USAID/Rwanda Valuing Open and Inclusive Civic Engagement Project.

[40] Wiley, J.,(1996). An Introduction to Categorical Data Analysis.