International Journal of Mathematics and Physical Sciences Research ISSN 2348-5736 (Online)
Vol. 5, Issue 1, pp: (41-54), Month: April - September 2017, Available at: www.researchpublish.com

Modeling the Determinants Influencing Labour Force Participation in Rwanda 2013/14, Using Logistic Regression Approach

[1]Gilbert MUGENZI, [2]Dr. Joseph K. Mung'atu, [3]Dr. Marcel Ndengo

[1]Student at (Master of Science in applied statistics), Jomo Kenyata University of Agriculture and Technology (JKUAT/KIGALI CAMPUS),
[2]Lecturer at Jomo Kenyata University of Agriculture and Technology/Nairobi, Kenya
[3]Lecturer at Jomo Kenyata University of Agriculture and Technology/ Kigali, Rwanda

Abstract: **This study established the influence of the determinants of labour force participation in Rwanda. It evaluated the roles and contributions of labour force participation in Rwanda, measured job creation for economic growth in Rwanda by using different policies like attracting foreign investors and modelling unemployment. The percentage of job seekers was high in youth and it was at 40% among Rwandan population. Workforce-skills matched what employers are looking for in order to measure the levels and trends of its indicators. The data used was secondary data from integrated household living condition survey (EICV4) for analysis, and this survey carried out by National Institute of Statistics of Rwanda and had sufficient data to avail labour force participation in Rwanda. Other findings from the study included how to increase the maximum job creation in Rwanda. It also used other secondary data from surveys and censuses carried out by national institute of statistics of Rwanda, like Fourth Rwanda population and housing census (4RPHC), establishment census, integrated business survey and labour force survey. The secondary data compiled was tested using Logistic regression model and other different tests like F- Statistic test and Chi-test analysis supported by fundamental evidence from the database of the regulatory authorities. Finally, it provided the model for solving problems identified by determinants.**

Keywords: **labour force participation, Employment, Unemployment, potential labour force, Self-employment.**

I. INTRODUCTION

1.1 Statement of the problem:

Labour force participation rate is used as an indicator to measure the economy and development of any country. The numbers of International Labour Organization [3] shows that labour force participation rate varies significantly from one country to another; it is an important characteristic for growth of a given country.

Unemployment affects not only job seekers, but also their families and the community in general but also the economy of the country as there could be many dependants compared to independent population size.

Therefore we need to analyze the determinants of labour force participation rate and its trend since [2] such that this can show a clear picture of employment and unemployment rate. This will inform policy makers some factors to consider in job creation and to some extent labour force skill mismatch.

The unemployment rate increases significantly with the job seekers. The unemployment rate of reproductive age also brings significant economic losses and setbacks to community development. Actual labour force participation is known/unknown as a main problem reason why I decide to do this research for clarifying all the factors influencing the increasing of unemployment rate and those problems hidden behind the ignorance, miss of information and provide the advices to policy makers in order to overcame in such situation of problems by establishing the policies for fighting against miss understanding of unemployment rate in Rwanda as negative impact of economic loss.

1.2 Objectives:

This research has general and specific objectives.

1.2.1 General Objective:

The general objective of this study is to model the determinants influencing labour force participation in Rwanda.

1.2.2 Specific Objectives:

Specific objectives assigned to this study are the following:

1. To determine the factors of labour force participation during the period of 2013-2014.
2. To measure the levels and trends of labour force participation rate in Rwanda.
3. To model the employment and unemployed by population characteristics in relation to labour force participation.

1.3 Research Questions:

1. What are the main determinants of labour force participation in Rwanda?
2. What are the levels and trends for labour force participation indicators in Rwanda?
3. What are the descriptions models of unemployment by background characteristics in Rwanda?

1.4 Justification of the study:

This study first of all was design in the purpose of obtaining a Degree in Master Science in Applied Statistics and it will be used by different people such are:

It can help the government to know, the factors influencing labour force participation in Rwanda as the mostly used determinants, the result find in data analysis will provide the advices for preventing and reducing unemployment rate in Rwanda.

It can help policy makers to know which sector they can improve according to the labour force participation in Rwanda. The result of this study will be published and thus become source of information helping to establish the policies for fighting against unemployment rate in Rwanda.

It can help other researchers to understand the methodology used in Rwanda to compute the labour force participation in Rwanda.

1.5 Scope of the study:

This research project was used the secondary data from the Integrated Household Living Conditions Survey (EICV) and Census (RPHC) to investigate the level of labour force participation in Rwanda. The data were gathered by visiting households on different period in order to aid household research covers these ten years, from 2002 to 2014, this study deals with determinants of labour force to capture the information related to the determinants influencing labour force participation in Rwanda. EICV and RPHC collects large and diverse information on labour force participation, however the effective and efficient implementation of both the data collection and the data processing some outliers and missing data remain untracked. In general the measurement of economic activity through the labour force participation in Rwanda is limited to a few indicators like employment and unemployment characteristics. It is also worth noting that the information presented in this research is limited to the main activity performed during the period of study, while the working population of Rwanda routinely works in multiple jobs.

II. METHODOLOGY

2.1 Sampling frame:

The sampling frame for the EICV4 and RPHC4 was based on a database of villages (umudugudu) that cover all of the households in Rwanda. This database includes information on all the geographic codes and the approximate number of households in each village. The geographic hierarchy of the villages in the sampling frame was based on the new administrative divisions of Rwanda: 5 provinces, 30 districts, 416 sectors, 2,148 cellules and 14,837 villages. The average number of households per village was 132 (168 for urban villages, 129 for rural villages). The urban and rural classification was based on the 2002 Rwanda Census of Population. In each sample village all the households were listed. This provided an updated sampling frame for the second stage of selection.

2.2 Sampling techniques:

Practicability of statistical sampling techniques allows the researchers to estimate the possible number of subjects that can be included in the sample, the type of sampling technique, the duration of the study, the number of materials, ethical concerns, availability of the subjects/samples, the need for the study and the amount of workforce that the study demands [1].

2.2.1 Stratification:

A stratified random sample is one obtained by separating the population elements into no overlapping groups, called strata, and then selecting a simple random sample from each stratum. See Richard L [6]. Regarding to the design of the study we found that the stratification sampling technique is the best sampling technique compare to other methods because the estimates for the survey were estimated up to district level considered as strata. "According to [6] they mentioned three reason for using stratification sampling technique and the nature of the study satisfy those conditions.

1. Stratification may produce a smaller bound on the error of estimation than would be produced by a simple random sample of the same size. This result is particularly true if measurements within strata are homogeneous.

2. The cost per observation in the survey may be reduced by stratification of the population elements into convenient groupings.

3. Estimates of population parameters may be desired for subgroups of the population. These subgroups should then be identifiable strata. The sampling frame of villages was stratified by district. Within each district the villages were ordered by urban and rural classification, then by geographic codes to provide an implicit stratification by urban and rural classification, and geographic location. This resulted in a proportional distribution of the sample villages by urban and rural classification.

2.2.2 Sample Size and Allocation:

The sample size was designed for measuring the Determinants influencing labour force participation in Rwanda. The idea is to let log p(x) be a linear function of x, modification of log p it became

$$\log\left(\frac{p}{1-p}\right) \tag{1}$$

We can make this a linear function of x.

The model often used to study the association between a binary response and a set of explanatory variables is given by logistic regression analysis. In this model, the natural logarithm of the odds ratio is related to the explanatory variables by a linear model. We will consider the situation where we have a single independent variable, but this model can be generalized to multiple independent variables. Let p(x) be the probability that y equals 1 when the independent variable equals x. We model the log-odds ratio to a linear model in x. [7]

$$\log\left[\frac{p(x)}{1-p(x)}\right] = \beta_0 + \beta_1 x_i + \beta_2 x_2 + ... + \beta_k x_k + \varepsilon_i \tag{2}$$

This transformation can be formulated directly in terms of p(x) as:

$$p(x) = \frac{e^x}{1+e^x} \tag{3}$$

And

$$p(x) = \frac{e^x}{1+e^x} = \frac{1}{1+e^{-x}} \tag{4}$$

The values of β_0 and β_1 would be estimated from the observed data using maximum likelihood estimation. We can interpret the parameters β_0 and β_1 in the logistic regression model in terms of p(x). The intercept parameter β_0 permits the estimation of the probability of the event associated with y=1 when the independent variable x=0.

Notice that the over-all specification is a lot easier to understand in terms of the transformed probability that in terms of the untransformed probability.1 To minimize the mis-classification rate, we should predict Y = 1 when $p \geq 0.5$ and Y = 0 when $p < 0.5$. This means guessing 1 whenever $\beta 0 + x \cdot \beta$ is non-negative, and 0 otherwise.

If we generalized linear models were formulated by John Nelder and Robert Wedderburn as a way of unifying various other statistical models, including linear regression, logistic regression and Poisson regression. They proposed an iteratively reweighted least squares method for maximum likelihood estimation of the model parameters [8].

$$E(y) = \mu = g^{-1}(x\beta) \tag{5}$$

: Expected value of $y = X\beta$ (6)

is the linear predictor g : Is the link function. The link function provides the relationship between the linear predictor and the mean of the distribution function.

β : Unknown parameter

$$Var(y) = V(\mu) = V(g^{-1}(x\beta)) \tag{7}$$

: Variance of the mean and link function will be equal to below equation.

$$y = X\beta = \frac{1}{1 + \exp(-x\beta)} \tag{8}$$

This shows the transformation of logistic regression model to the linear regression model and logistic regression model tested by χ^2 .

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_2 + ... + \beta_k x_k + \varepsilon_i \tag{9}$$

Distribution (%) of usual economic activity status, according to urban/rural, province, sex and four-year age group (EICV4)

This method helps in data presentation, curve presentation; table's construction and figures interpretation of the logistic regression model.

The success prediction model that will be used in this study will use five variables that have significant determinants of labour force participation rate. The variables that will be used in this study are grouped as follows:

1. Dependant variable
2. Independents variables

2.3 Model:

Given the nature of variables in this study a logistic regression model used to statistically measure the level of relationship between dependent variables and independents variables of labour force participation in Rwanda. Logistic regression measures the relationship between the categorical dependent variable and one or more independent variables by estimating probabilities using a logistic function, which is the cumulative logistic distribution. Logistic regression can be seen as a special case of the generalized linear model and thus analogous to linear regression. The model of logistic regression, however, is based on quite different assumptions (about the relationship between dependent and independent variables) from those of linear regression. In particular the key differences of these two models can be seen in the following two features of logistic regression. [4]. Logistic regression can be binomial, ordinal or multinomial. Binomial or binary logistic regression deals with situations in which the observed outcome for a dependent can have only two possible types (dead vs. alive or win vs. loss).

In binary logistic regression, the outcome is usually coded as 0 or 1, as this leads to the most straightforward interpretation. If a particular observed outcome for the dependent variable is the noteworthy possible outcome (referred to as a success or a case) it is usually coded as 1 and the contrary outcome (referred to as a failure or a non case) as 0. Logistic regression is used to predict the odds of being a case based on the values of the independent variables (predictors). The odds are defined as the probability that a particular outcome is a case divided by the probability that it is a non case [5].

As a predictive analysis, logistic regression model describes data and explains the relationship between one dependent variable and two or more independent variables. In this analysis, the dependent variable is labour force participation in Rwanda (employment and unemployment rate) and the independent variables are:

1. Age structure (16+ in age group)
2. Sex (male, female)
3. Education(none, primary, secondary and University)
4. Area(urban, rural) of residence
5. Marital status(single, married, divorced and widow/widower)
6. Province

2.3.1 Estimation of parameters:

In the model mentioned above, parameters will be interpreted as follows:

1. When relationship exists between two variables the correlation coefficient is positive or negative.
2. When there is variation in one variable related to variation to other variable we call it the coefficient of determination.

2.3.2 Overall model fit:

The most common assessment of overall model fit in logistic regression models is the goodness of fit test.

The goodness of fit of the model will be found after running Chi Square test**,** if the P -value of the parameter is below 0.05 we will conclude the model is statistically significant and it will prove that the independent variables have a significant impact on the dependent variable. How well the model fits the observed data is assessed by a number of ways as follows: A variety of statistical tests [4] are applied with their P- values determined in order to assess how well the model describes the data. With grouped data we can assess goodness of fit by looking directly at the deviance, which has approximately a chi-squared distribution for large n_i . A common rule of thumb is to require all expected frequencies (both expected successes $\hat{\mu}$ and failures $(n_i - \hat{\mu}_i)$) (10)

We use chi-square as a measure of model fit here in a similar way. It is the fit of the observed values (Y) to the expected values (Yˆ). The bigger the difference (or deviance) of the observed values from the expected values, the poorer the fit of the model. So, we want a small deviance if possible. As we add more variables to the equation the deviance should get smaller, indicating an improvement in fit.

2.3.3 Chi-Square –test:

Chi-square is a statistical test commonly used to compare observed data with data we would expect to obtain according to a specific hypothesis. Measure a single categorical variable on each person or trial. Each person or trial falls into one of k mutually exclusive categories. Null hypothesis specifies the probabilities of falling into each of the k categories. Alternative hypothesis is that those are not all correct

Step 1: Determine the null and alternative hypotheses.

H0: The probabilities for k categories are p1, p2. . . pk. H0: Oi = Ei (11)

Ha: Not all probabilities specified in H0 are correct. Ha: Oi ≠ Ei (12)

Step 2: Verify necessary data conditions, and if met, summarize the data into an appropriate test statistic. Data condition needed: At least 80% of the expected counts are greater than 5 and none are less than 1. Hypothesis, the test statistic is:

$$\chi^2 = \sum \frac{(Observed - Expected)^2}{Expected} \quad (13)$$

Where the expected count for the ith category is computed as npi.

Step 3: P-value of Chi-square Test Large test statistic tend evidence that values in null are not correct (observed counts don't match expected counts). P-value equal to probability the chi-square test statistic could have been as large or larger if the null hypothesis were true. Chi-square probability distribution used to find p-value. Degrees of freedom: df = k – 1 (14)

Step 4: Making a decision. Large test statistic tends to small p-value. Evidence that the proportions are not as specified. Two equivalent rules:

1. Reject H0 when Ha is significant.

2. 0.05≤ p-value

Chi-square statistic is greater than the entry in the 0.05 (the critical value). That defines the rejection region.

Step 5: Conclusion: Pennsylvania lottery digits drawn are not statistically different from what's expected by chance.

The distribution of the quantity χ^2 can be approximated by a chi-square distribution provided that the expected value Ei are fairly large. The chi-square goodness-of-fit test based on k specified cell probabilities will have k -1 degrees of freedom. The approximation of the sampling distribution of the chi-square goodness-of-fit test statistic by a chi-square distribution improves as the sample size n becomes larger.

2.4 Data collection:

Data collection of EICV4 took 12 months which is divided in ten cycles. Each cycle covered 33 days in Kigali city province where enumerators visited a household after two days and 16 days in other provinces and enumerators visited the household after one day. In Kigali city province in each cycle the enumerators covered 5 clusters and in other provinces they covered 2 clusters. During the period of data collection a calendar was drafted and in each day of visit the enumerator has the section to be asked.

2.5 Data processing:

Data entry of the completed and checked questionnaires was undertaken at National Institute of statistics of Rwanda (NISR) offices by trained staff members using CSPro software. Thereafter, data was entered in computers, exported in SPSS software and edited. The resulted in the introduction of CSPro batch editing; use of SPSS for editing and computing range checks (transformed to look up tables later used by CSPro).This dataset are available at NISR website (www.statistics.gov.rw) where people can access it for different analysis.The tools used in data analysis are Statistical Software IBM 21 version and STATA vision 13 were used in data analysis and interpretation were done through the use of test of deviation, p- value, analysis of standard deviation and Chi-Square test .Once calculated, the chi square statistic is compared to a critical value, which is determined using the sample size, the chosen level of significance (typically a=0.05), and a critical value in the table. If the statistic calculated is less than the critical value, the sample fits the theoretical distribution. Otherwise, there is a significant different between the observed and expected.

III. RESEARCH FINDINGS AND DISCUSSION

3.1 Distribution of the outcome variable:

The analysis start with first explain the method which used for analyzing the data at hand, then after finding the model, The data obtained from household questionnaires relating to Integrated Household Living condition Survey [2] Data collected at the individual, household and community level, therefore allowing for individual-level independence of socio-economic characteristics. Given the labour market participation focus of this paper, we restrict our sample to working aged adults.

This research decided the test like Chi-square which statistical test to be used in measuring the factors influencing labour force participation in Rwanda. There is type of statistical test namely-test. Logistic regression model, which might be undertaken depending on the data distribution. As the stratified random sampling techniques have been adopted for the survey, below are the formula used to estimate various statistical parameters. Some notation is required for stratified random sampling. The suffix h denotes the stratum and i the unit within the stratum.

The methodology includes the general approach of the study ,the population of interest, the sample, data collection instruments and the data analysis technique that were used.

Table 3.1: Labour force participation rate in Rwanda (Anova table)

Number of obs = 36980 R-squared = 0.2313
Root MSE = .305337 Adj R-squared = 0.2309

Source	Partial SS	df	MS	F	Prob > F
Model	1036.91428	22	47.1324675	505.55	0.0000
province	23.7009209	4	5.92523023	63.55	0.0000
ur2_2012	18.5458039	1	18.5458039	198.92	0.0000
sex	1.23182896	1	1.23182896	13.21	0.0003
Age_group	291.949494	10	29.1949494	313.15	0.0000
Mar_stat	23.926753	3	7.97558433	85.55	0.0000
edu_level	161.647877	3	53.8826257	577.95	0.0000
Residual	3445.52647	36957	.093230686		
Total	4482.44075	36979	.121215846		

Referring to results presented in Table 3.1 The ANOVA table above has shown how the explanatory variables explained the model and relationship between the variables. As we have different F-test we have to check every step of testing hypothesis.

Before we have to compute critical value in order to compare with observed one on the table,

Then, critical is F (22, 36957) =1.84

Step 1: H0: $\mu 1 = \mu 2 = \mu 3 = \mu 4$, i.e. treatments are equally effective

HA: The means are not all equal.

Step 2: An F statistic is appropriate, since the dependent variable is continuous and there are 2 or more groups.

Step 3: Since $\alpha = .05$ and df = (22, 36957), accept H0 if F1.84 > 2.57

Step 4: The computed value of the F statistic is 1.84

Step 5: Accept H0. The treatments are not equally effective so, we have to accept null hypothesis which is not significant, then we reject alternative one (Ha).

Also p-value is less than 0.05 we reject null hypothesis (H0) then we accept the alternative hypothesis (Ha).So, the null hypothesis is significant effect because the variables prove the relationship between them.

Table 3.2: Distribution of labour force participation Economic activity

Current economic activity status[All HH members 16+]	Freq.	Percent	Cum.
Currently unemployed	919,572	14.37	14.37
Currently employed	5,478,572	85.63	100
Total	6,398,144	100	

The distribution shows that the employed peoples are 86% of population aged 16 and above then the unemployed one equal to 14%.This increments depend on different causes or factors as it was said before.

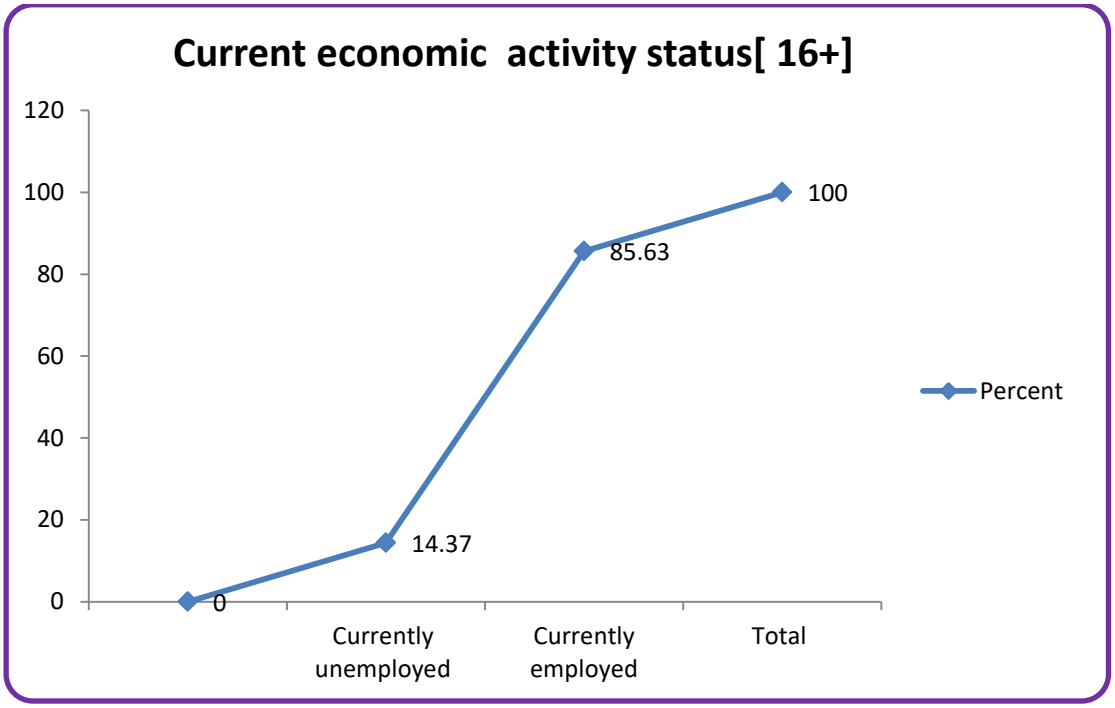

Figure 3.1: Current economic activity

3.2: Independent variables:

Table 3.3: Distribution of labour force participation by Age group

Age Group Categorization	Freq.	Percent	Cum.
16-19Yrs	961,225	15.02	15.02
20-24Yrs	1,080,232	16.88	31.9
25-29Yrs	909,845	14.22	46.12
30-34Yrs	841,710	13.15	59.27
35-39Yrs	557,088	8.7	67.97
40-44Yrs	445,830	6.97	74.94
45-49Yrs	357,274	5.58	80.52
50-54Yrs	361,498	5.65	86.17
55-59Yrs	297,158	4.64	90.82
60-64Yrs	203,978	3.19	94
65+Yrs	383,810	6	100
Total	**6,399,650**	**100**	

The distribution of age shows that the high percentage of age group are in 20-24 are around 17% of years and the following one are in age group 16-19 are 15% also 25-29 have 14.22% and 30-34 years which is around 13.15% , this means that more than 50% of working age are youth or under 35 years old, then the lowest percentage of age group is in group 60-64 years which it has 3.19% of people having age of working in Rwanda and this also has an effect on labour force because when age increasing then the number or percentage peoples in labour force decrease which have a negative impact in the society because of lack of experienced persons. So, every additional of year decline the chance of having a job or increase the probability of missing a job.

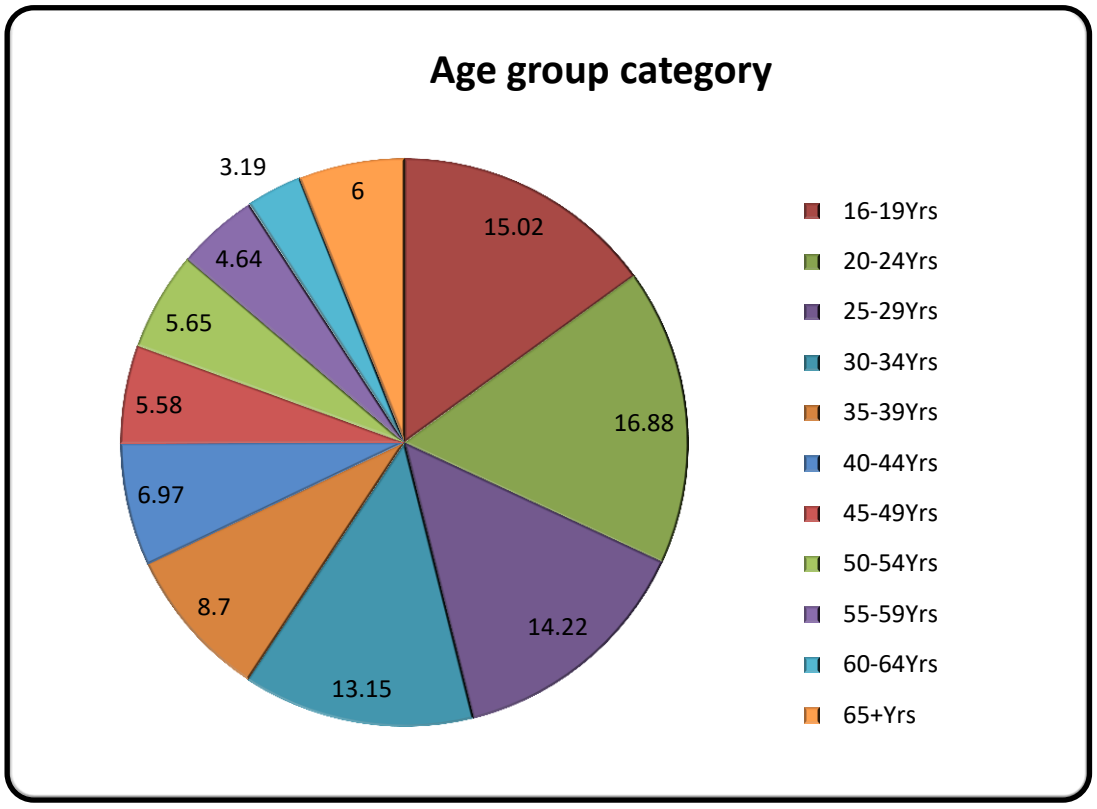

Figure 3.2: Distribution of labour force by age group

This figure shows the distribution of age group in labour force participation and each group category has their own meaning especially age group 20-24 has 16.88% of all peoples who has age of working followed by 16-19 with 15.02% and lastly is old peoples find in age group 60-64 with 3.19% as small numbers of labour force participation.

Table 3.4: Distribution of labour force participation by Sex

Sex of Respondent	**Freq.**	**Percent**	**Cum.**
Male	2,970,114	46.41	46.41
Female	3,429,536	53.59	100
Total	**6,399,650**	**100**	

The distribution shows that the share of females is high compared to a male which is around 54% of population aged 16+ and male have 46% of peoples having age of working.

Table 3.5: Distribution of labour force participation by marital status

Marital Status	**Freq.**	**Percent**	**Cum.**
Married	3,451,944	53.94	53.94
Divorced	179,652.27	2.81	56.75
Single	2,251,416	35.18	91.93
Widow/widower	516,637.22	8.07	100
Total	**6,399,650**	**100**	

This table shows that the married group had highest percentage with 53.9% and the second one is single group whom had 35% of population aged 16 and above. The last one is population of divorced with around 3%.

Table 3.6: Distribution of labour force participation by Education

Education level of completed formal education	Freq.	Percent	Cum.
Never attended	1,211,106	18.93	18.93
Primary	4,240,208	66.26	85.18
Secondary	796,608	12.45	97.63
University	151,548.99	2.37	100
Total	**6,399,472**	**100**	

This table shows the distribution between the level of education and their variation. So, as the age of working increase the educated persons decrease. In primary level the percentage is high with 66%, never attended around 19%, secondary12.4% employed peoples compared with the university holders which has 2% of employment.

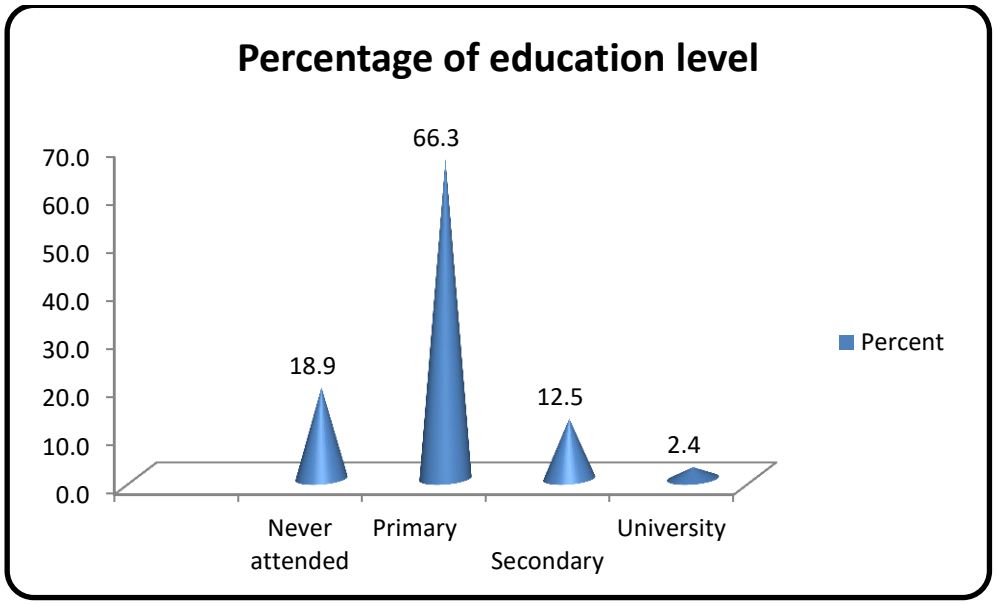

Figure 3.3: Distribution of education attainment in labour force

This figure above shows the share of different level of education by attainment in labour force participation in Rwanda. The educational attainment of the youth population aged 16 to 30 years old in Rwanda, the labour force participation tends to be higher than the overall labour force participation. About 66.3% of employed persons has primary education and never attended has 18.93% of labour force as an employed peoples in Rwanda. Finally educated persons with university level the percentages decries to 2.4% of employed peoples. In terms of gender, educational attainment among the employed persons female in the labour force tends to be high compared to male with 46%. Lastly educated peoples with university level 2.4% are less likely being employed compared to the persons with primary level. So, this means that micro establishments offer more jobs for those whom who have low education or technical skills.

Table 3.7: Distribution of labour force participation by Area (urban/rural)

Area of residence	Freq.	Percent	Cum.
Rural	5,182,724	80.98	80.98
Urban	1,216,926	19.02	100
Total	**6,399,650**	**100**	

Around 81% of the Rwandan population lives in rural areas, and 19% living in urban areas at the national level. This means that the higher population living in rural area because of different factors will be proved in the model of this research.

Table 3.8: Distribution of labour force participation by Province

Province	Freq.	Percent	Cum.
Kigali City	821,978	12.84	12.84
Southern Province	1,501,672	23.46	36.31
Western Province	1,441,057	22.52	58.83
Northern Province	1,010,809	15.79	74.62
Eastern Province	1,624,135	25.38	100
Total	**6,399,650**	**100**	

In all provinces, the majority of population lives in rural areas, except in Kigali City where percentage are 12.84% of peoples living in urban of its total population. Southern Province had the first highest percentage of peoples living in rural areas which is 23.4% and Western Province has the second high percentage of the population living in rural areas which is 22.52% of total population. The last one of the province which has the lowest percentage of living in rural is Northern with around 16%.

Table 3.9: Parameter Estimates of the logistic regression model

Survey: Logistic regression

Number of strata = 30 Number of obs = 36980
Number of PSUs = 1230 Population size = 6397965.3
Design df = 1200
F(22, 1179) = 203.64
Prob > F = 0.0000

employment	Linearized Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Age_group						
20-24Yrs	1.317695	.0529997	24.86	0.000	1.213713	1.421678
25-29Yrs	2.237796	.0762522	29.35	0.000	2.088193	2.387398
30-34Yrs	2.547339	.1122242	22.70	0.000	2.327162	2.767516
35-39Yrs	2.656804	.1304895	20.36	0.000	2.400792	2.912817
40-44Yrs	2.217362	.1458005	15.21	0.000	1.931309	2.503414
45-49Yrs	2.082042	.1450303	14.36	0.000	1.797501	2.366583
50-54Yrs	2.032021	.1509737	13.46	0.000	1.735819	2.328223
55-59Yrs	1.347795	.1296631	10.39	0.000	1.093403	1.602186
60-64Yrs	.9139046	.1433621	6.37	0.000	.6326363	1.195173
65+Yrs	-.2910427	.1041238	-2.80	0.005	-.4953276	-.0867578
Mar_stat						
Married	.175201	.0770788	2.27	0.023	.0239768	.3264253
Divorced	.6572088	.2001163	3.28	0.001	.2645921	1.049826
Single	-.9180322	.1012694	-9.07	0.000	-1.116717	-.7193475
sex	-.1705701	.0390163	-4.37	0.000	-.2471178	-.0940224
Province						
Kigali City	-.5070208	.0936405	-5.41	0.000	-.6907381	-.3233034
Western Province	.2301437	.0685237	3.36	0.001	.095704	.3645833
Northern Province	.4879767	.0748384	6.52	0.000	.3411481	.6348053
Eastern Province	.069725	.0618116	1.13	0.260	-.0515458	.1909959
edu_level						
Primary	.0968363	.0645525	1.50	0.134	-.029812	.2234846
Secondary/Post-primary	-1.518475	.0733981	-20.69	0.000	-1.662478	-1.374472
University	-1.2859	.1222349	-10.52	0.000	-1.525718	-1.046082
Urb_rural	-.6530953	.0783114	-8.34	0.000	-.8067378	-.4994528
_cons	1.487569	.1307828	11.37	0.000	1.23098	1.744157

3.3 Interpretation of the results in terms of coefficients of the logistic regression model:

By interpreting the coefficients of the Logistic regression model, the positive coefficients indicate that there is an increment of employment influenced by factor, while the negative coefficients indicates that there is also an increment of unemployment influenced by determinants of labour force participation.

$$y_{ij} = 1.48 + 1.31x_{10} + 2.23x_{11} + 2.54x_{12} + 2.65x_{13} + 2.21x_{14} + 2.08x_{15} + 2.03x_{16} + 1.34x_{17} + 0.91x_{18} - 0.29x_{19} - 0.17x_{20} + 0.09x_{31} - 1.51x_{32} - 1.28x_{33} - 0.65x_{41} + 0.17x_{51} + 0.65x_{52} - 0.91x_{53} - 0.50x_{61} + 0.23x_{63} + 0.48x_{64} + 0.06x_{65} \quad (15)$$

The variable age as it is categorical also shows that in age group 60-64 years are more likely to have a job compared to the others working age group with 91% followed by age group35-39 with 65% and 30-34 with 54% , but age group16-19 is omitted and 65 and above are less likely to be employed with -29%. So, youth and old peoples are the two groups who have less chance of being employed which means that 65 and above should be removed even if the p-value is less than 5% which is significant. Also age group is symbolized by x_{1i} .

The total labor force size that was captured in the Rwanda EICV4 study is 36,980 representing 82% of total respondents. The gender distribution of the respondents shows that 46.4% are male while the remaining 53.5% are female. This means that gender disaggregation shows that female had lower chance of being employed because their percentages are -17% compared to male with 46% which is statistically significant. And sex as variable is replaced by x_{2i} in the model.

This model shows that more you have higher level of education more you have a big chance of not being employed also if you have low level of education more you have big chance of being employed. This means that the percentage of being employed with secondary level of education is equal to -51% compared to university level which is equal to -28%,then after the percentage of being employed with primary level of education is equal to 9% which is high compared to the others level of education. So, the low level of education is accepted as null hypothesis but statistically not significant because their p-value is greater than 0.05. The level of education variable is replaced by x_{3i} .

Also the model shows that the peoples living in rural area increasing the chance of being employed compared to the one who live in the urban area, because the percentages of being employed with rural area is 35%, then in urban is -65%.The implication of area of residence is that living in urban area imply not having the chance of being employed. Also area of residence is symbolized by x_{4i} .

As others variables shows effect on employment also marital status it has. The divorced peoples had percentages of being employed which is equal to 65% and followed by married one their percentages of having a job is equal to 17% then the single populations are more likely not accessed to the job because their percentages are -91%. So, the single persons should be removed on labour market in Rwanda even if their p-value is less than 5% which is significant, and marital status is symbolized by x_{5i} .

On province level replaced by x_{6i} peoples who lives in Northern Province had high percentages of being employed with 48%, followed by Western province with 23% and the Eastern province had the low percentages with 6% compared to the others provinces. But someone who lives in Kigali city he/she is more likely not being employed because it decreases linearly with -5%. So, it means that the null hypothesis is statistically significant because p-value less than 5% accept in Eastern province with 0.26 p-value which will be accepted null hypothesis and not significant.

IV. CONCLUSION AND RECOMMENDATION

4.1 Conclusion:

This research has been analyzed the main socio-economic characteristics associated with labour force participation rate during a data collection of EICV4 in the module of different explanatory variables. In the analysis, conditional on labour force participation rate, the research has analyzed the role played various factors (employment and unemployment characteristics) in influencing the labour force participation rate into different areas of market.

Having analyzed the logistic regression model coefficients, it is evident that sex ratio plays a great role with 35% of male and increasing labour force participation rate for rural area with 83.9% and urban area with 13%.So, it is better to live in rural are based on research proof.

It was possible to analyze interprovincial variation because living in northern means to have a chance of being employed on 48% and because the coefficients that behaved differently across rural and urban areas are statistically significant with p-value which is less than 5%. A pooling across rural-urban data allows us to examine the effect of explanatory variables on the dependant variable for Rwanda. This analysis has been carried out for the variables that have been studied. These omitted variables have been also replaced by constant in the model.

The household data analyzed in this research has several important findings. First, as expected, education has a strong negative impact on labour force participation. Using the categorical variables for years of schooling, our estimates confirm that an additional year of schooling decreases the chance of being employed because a person who has university level are less likely being employed. Indeed, the more we include lower levels of education in Rwanda more you increase the chance of having a job. However, breaking the effect down by the three levels of education, our results emphasize the important influence of having university or secondary education and primary level in labor force participation has negative impact.

Also married peoples living in rural area with age group between 25-29 living in Eastern province are more likely to be employed compared to one who is single living in urban area in Kigali city with age group 20-24 and with university level and being female are more likely to be unemployed

4.2 Recommendations:

Economic policies which control the number of unemployed and employed peoples per family like Rwanda's where Policy should be strictly enforced. This would gradually improve the ability of peoples to join the labour force. Controlling the population in this sense can also solve other economic issues that cripple this huge economy such as poverty, economic inequality, loss of hope etc. This can also help minimize the probably negative effect of having big numbers of youth unemployed in the household.

Proportion of agricultural households could be an important determinant of labour force participation rate in rural because the most population living in rural area are more likely to have a job. Hence, the government should encourage youth in rural area to participate in agricultural sector as economic activities also engage in non-agricultural jobs/tasks by providing incentives such as plots per age group. This can also improve the effect of educated persons like youth whom they have around a half of population who had age of working to focus on technical training instead of formal education because it gives more jobs.

The government of Rwanda should take the policy of promoting self employment in different sectors specifically in the urban areas because from the skilled persons the innovation economy increase faster and it reduce the poverty in the population. Also educated peoples should improve agriculture sector again and it will generate more income for big population. Also the policy makers should put in place to facilitate youth into development by motivating them through gathering in cooperative or association for self/own employment. Other recommendation is that as youth is a part of big number of Rwandan population and most of them are single when youth will be facilitated to self-employment it will promote household and at the same time the economy of country.

As I mentioned above the policy makers should think on educated peoples and the quality of education because when the level of education increase the chance of being employed decrease, and institutions looking for employees at the same time. So, those issues it should be considered as the main one to solve because promoting education for all without occupation if a big burden to the country.

REFERENCES

[1] Bowen, W.G. and Finnegan, 1969.The Economics of Labour force participation, Princeton N.J.: Princeton university press, 1969.

[2] http://www.statistics.gov.rw/datasource/integrated-household-living-conditions-survey-4-eicv-4. Economic Activity.

[3] ILO (1982), "Labour Force, Employment, Unemployment and Underemployment", Report II, Thirteenth International Conference of Labour Statisticians, ICLS/13/11, Geneva.

[4] Hosmer, D. W. and Lemeshow, S. ÔA goodness-of-Þt test for the multiple logistic regression model, Communications in Statistics, A10, 1043-1069 (1980).

[5] Cox DR, Snell EJ. Analysis of Binary Data. 2nd edn. Chapman and Hall: London, 1989.

[6] Richard L. Scheaffer, William Mendenhall III, R. Lyman Ott and Kenneth Gerow .1996. Elementary survey sampling,Seven Edition

[7] McCullagh, Peter; Nelder, John (1989) Generalized Linear Models, Second Edition. Boca Raton: Chapman and Hall/CRC. ISBN 0-412-31760-5.

[8] Nelder, John; Wedderburn, Robert (1972). "Generalized Linear Models". Journal of the Royal Statistical Society. Series A (General). 135 (3): 370–384. doi:10.2307/2344614. JSTOR 2344614