International Journal of Mathematics and Physical Sciences Research ISSN 2348-5736 (Online)
Vol. 3, Issue 1, pp: (134-149), Month: April 2015 - September 2015, Available at: www.researchpublish.com

MODELING FACTORS ASSOCIATED WITH INTIMATE PARTNER VIOLENCE AGAINST WOMEN IN INTIMATE PARTNERSHIP IN RWANDA

[1]Elysee TUYISHIME, [2]Dr. Joseph K. Mung'atu, [3]Dr. Jairu NDIGA DESIDERIO

[1]Students at Jomo Kenyata University of Agriculture and Technology (JKUAT/KIGALI CAMPUS),
Master of Science in applied statistics
[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*: Intimate partner violence (IPV) against women is a global public health and human rights concern in Rwanda. As reviewed from two national surveys, in Rwanda, women's experience of physical or sexual IPV in their lifetime almost doubled from 34% in the 2005 DHS, to 56% in the 2010 DHS, while globally IPV rates at 30% and in other sub-Saharan African countries, such as Uganda (40%); Zimbabwe (37%) and Kenya (24), placing Rwanda among the countries with the highest rates of IPV against women in the world, in order to inform the design of IPV prevention programs. A secondary analysis of RDHS 2010 was done and descriptive statistics was used to summarize continuous data, and categorical data, bivariate statistics was used to compare two study groups to provide strong evidence of any study group differences, The test for collinearity was done before fitting Multivariate logistic regression model to produce parsimonious (efficient) multivariable models, backward stepwise logistic regression was used to find the final model. Of the 3042 women included in the analysis 56.6% (n=1718) had experienced at least one form of IPV in the last 12 months prior to the survey. Women corresponding to 51.9% (n=1580) reported to have experienced severe form of IPV whereas other women equivalent to 33% (n=1007) reported a less severe form of IPV. Young women whose husband have not attained high school and having multiple sex partners, living in a high density household in rural areas are more likely to face IPV. IPV prevention programs should increase focus development initiatives to improve access to education for girls and boys may also have an important role in violence prevention.**

***Keywords:* Intimate Partner Violence, Simple Logistic Regression, Multiple Logistic Regression.**

I. INTRODUCTION

1.1 Statement of the problem:

Intimate partner violence (IPV) towards women is an important issue to address to since it affects family's welfare economically and socially as well; and for the country as a whole, however often neglected public health issue. The existence of gender norms imbalance expressed by men's and women's attitudes in relation to power and decision-making in intimate relationships may influence the degree of IPV. The aim is to examine potential risk factors of physical, sexual and psychological IPV among couples in Rwanda.

The Government of Rwanda inspired by the philosophy behind the Convention on the Elimination of All Forms of Discrimination and Violence against Women to achieve the Millennium Development Goals (MDGs. In this context, there is a need of evidence-based policies on the country level to address this serious(GMO 2013).

As Rwanda is classified among the highest levels of IPV in the world, therefore, there is still a gap to be filled and this needs an insights into the problem in order to empower policy makers to understand violence against women, especially preventing this type of violence. Also based on the recommendations from this research appropriate interventions will hopefully be set with the purpose of preventing IPV in Rwanda, it is with this need that this project has been put forward.

1.2. Objectives:

1.2.1General Objective

To find the prevalence of IPV and model potential factors associated with IPV against women in intimate partnership in Rwanda.

1.2.2Specific Objectives

This study had the following Specific Objectives:

1. To determine the extent of IPV in Rwanda
2. To identify potential demographic and socioeconomic factors associated with IPV in Rwanda.
3. To model potential risk factors associated with IPV in Rwanda.

1.3. Research Question:

What factors are associated with the extent of all forms of Intimate Partner Violence against women in intimate partnership in Rwanda?

1.4. Justification:

Rwanda is classified among the highest levels of IPV in the world, with national estimates showing that 55.6 percent of women have experienced physical violence and 17.5 percent have experienced sexual violence in the past 12 months from their current or most recent husband/partner (Mannell & Jackson 2014). Therefore, if we can be able to identify potential risk factors associated with IPV, it will be possible for policy maker to understand violence against women, including preventing this type of violence and helping women to recover. Findings from this study will be published so that interested organizations and institutions be informed ways to act early to prevent the occurrence of intimate partner violence and promote health in Rwanda.

Prevention approaches needs evidences to be based on to understand potential risk factors and respond to the intimate partner violence public health problem. This research aims at providing a mathematical model that will describe the extent at which individual risk factors influence the IPV.

II. METHODOLOGY

2.1 Sampling technique:

2.1.1 Selection probability and Sampling weight

The sampling frame for the Integrated household living condition survey (EICV3) 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, 2148 cellules and 14837 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.

Because of the non-proportional allocation of the sample to the different provinces and to their districts and the possible differences in response rates, sampling weights is required for any analysis using 2010 RDHS data; this ensures the actual representativeness of the survey results at the national level as well as at the domain level. Because the 2010 RDHS sample is a two-stage stratified cluster sample, sampling weights was calculated based on separate sampling probabilities for each sampling stage and for each cluster. The Following notations were used:

P_{1hi} : First-stage sampling probability of the ith village in stratum h

P_{2hi} : Second -stage sampling probability within the ith village (household selection)

Let a_h be the number of villages selected in stratum h, M_{hi} be the total population according to the sampling frame in the ith village and $\sum M_{hi}$ be the total population in the stratum h. The probability of selecting the ith village in the 2010 RDHS sample is calculated as follows:

$$\frac{a_h M_{hi}}{\sum M_{hi}} \quad (1)$$

Let b_{hi} be the proportion of households in the selected segment compared with the total number of households in the village i in stratum h if the village is segmented; otherwise $b_{hi} =$ 1. Then the probability of selecting village i in the sample is:

$$P_{1hi} = \frac{a_h M_{hi}}{\sum M_{hi}} * b_{hi} \quad (2)$$

A 2010 RDHS cluster is either a village or a segment of a large village. Let L_{hi} be the number of households listed in the household listing operation in the cluster i in stratum h, Let g_{hi} be the number of households selected in the cluster. The second stage's selection probability for each household in the cluster is calculated as follows:

$$P_{2hi} = \frac{g_{hi}}{L_{hi}} \quad (3)$$

The overall selection probability of each household in cluster i of stratum h is therefore the production of the two stages of selection probabilities:

$$P_{hi} = P_{1hi} * P_{2hi} \quad (4)$$

The design weight for each household in cluster i of stratum his the inverse of its overall selection probability:

$$W_{hi} = \frac{1}{P_{hi}} \quad (5)$$

The next is design weights, design weights was adjusted for household nonresponse as well as for individual non-response to get the sampling weights for women's and men's surveys, respectively. The differences in the household sampling weights and the individual sampling weights are introduced by individual nonresponse. The final sampling weights was normalized to give the total number of unweighted cases, equal to the total number of weighted cases at the national level, for both household weights and individual weights, respectively. The normalized weights are relative weights, which are valid for estimating means, proportions, and ratios.

2.1.2. Controlling for complex survey design

As the DHS methodology is complex, a systematic approach is needed for this secondary analysis to construct a prepared dataset for analysis. This analysis will consider three dataset to be combined in order to get full information needed for analysis. The three datasets will be merged including; Individual women's recode (IR) file, Male recode (MR) file, and Household recode (HR) file.

Several characteristics of complex survey design can bias mean and variance estimates. Any survey design characteristic which effects the probability of selection including stratification, oversampling, and response rates must be accounted for with the application of sampling probability weights in descriptive data analysis. Descriptive data analysis must also adjust for clustering by widen the variance estimates to avoid making type I errors. Accounting for stratification can slightly narrow confidence intervals in analyses of multiple strata, but the effect is usually negligible, and so the effect of stratification on variance can be ignored in descriptive data analysis(Balian et al. 2014).

2.1.3. *Descriptive analysis*

This analysis will cover three descriptive statistics; means and medians will be used to summarize continuous data, and percentages will be used to summarize categorical data.

2.1.4. Bivariate analysis

Bivariate statistics will be used to compare two study groups to see if they are similar. When comparing groups, we want to provide strong evidence of any group differences, so we will use a conservative threshold of $p<0.05$ to determine statistical significance. Since research questions is with binary outcomes, bivariate statistics will be used to summarize and compare characteristic across groups.

We will also use bivariate statistics to identify potential covariates that are worth testing in a multivariable model. If a variable is independently associated with the outcome, it might continue to explain the outcome once other factors are taken into account. In this case, when bivariate statistics are used for the purpose of filtering potential covariates in multivariate analysis, we use a generous threshold of p-value<0.1 to determine statistical significance to ensure that we do not drop any potentially useful variables from the analysis. Here the same statistical test that will be used to compare two groups is the chi-square test in logistic regression, is the same test and output that we will use here to filter variables will be the same. The only difference is in purpose of the test, and therefore our interpretation of its results will be different.

Pearson's chi-square test will be used to test whether the distribution in a categorical variable is statistically different in two or more groups. The chi-square test gives a yes/no answer a p-value less than the threshold will mean, yes, there are differences between the two groups.

A t-test will be used to test whether the distribution of a continuous variables are statistically different across groups, a p-value less than the threshold will mean, yes, there are differences

Before fitting any kind of multivariate model whether a general explanatory model or a hypothesis test model we will need to test for collinearity.

Collinearity occurs when two covariates in a multivariable model are highly related; usually this is because the two variables represent the same thing (the same concept, or they happen simultaneously). As a result, the model becomes unstable. To produce parsimonious (efficient) multivariable models, and to prevent strange, unstable results, we test for strong associations among covariates and remove any collinear covariates from the analysis.

The Pearson's R correlation coefficient will be used to identify binary, ordinal, and continuous covariates that are correlated. Correlations of $r>0.5$ will be considered as collinear as in the social sciences. When two or more covariates will be found to be collinear, we will keep the one variable that will be most strongly associated with the outcome, unless there will be a conceptual reason to keep one over the other.

2.1.4.1 Simple logistic regression model:

So we approach this problem by using linear regression;

1. the most obvious idea is to let p(x) be a linear function of x. Every increment of a component of x would add or subtract so much to the probability. The conceptual problem here is that p must be between 0 and 1, and linear functions are unbounded thus Linear models can't do this.

2. The next most obvious idea is to let log p(x) be a linear function of x, so that changing an input variable multiplies the probability by a fixed amount. The problem is that logarithms are unbounded in only one direction, and linear functions are not.

3. Finally, the easiest modification of log p which has an unbounded range is the Logistic (or logit) transformation, log

$$\frac{p}{1-p} \qquad (6)$$

We can make this a linear function of x without fear of nonsensical results. (Of course the results could still happen to be wrong, but they're not guaranteed to be wrong.) This last alternative is logistic regression(Project 2011).

Formally, the logistic regression model is that;

$$\text{Log}\frac{p(x)}{1-p(x)} = \beta 0 + x \cdot \beta 1 \qquad (7)$$

Solving for p, this gives

$$P(x;\ \beta_i) = \frac{e^{\beta_0 + x \cdot \beta_i}}{1+e^{\beta_0 + x \cdot \beta_i}} = \frac{1}{1+e^{-(\beta_0 + x \cdot \beta_i)}} \qquad (8)$$

2.1.5. Multivariate logistic regression analysis:

More generally, consider a random variable Z that can take on one of two possible values. Given a dataset with a total sample size of M, where each observation is independent, Z can be considered as a column vector of M binomial random variables Zi. By convention, a value of 1 is used to indicate "success" and a value of either 0 or 2 (but not both) is used to signify "failure." To simplify computational details of estimation, it is convenient to aggregate the data such that each row represents one distinct combination of values of the independent variables. These rows are often referred to as "populations." Let N represent the total number of populations and let n be a column vector with elements ni representing the number of observations in population i for i = 1 to N where $\sum_{i=1}^{N} n_i$ = M, the total sample size. Now, let Y be a column vector of length N where each elements Yi representing the number of successes of Z for population i. Let the column vector y contain elements yi representing the observed counts of the number of successes for each population. Let π be a column vector also of length N with elements πi=P (Zi=1/i), i.e... The probability of success for any given observation in the ith population(Pohar et al. 2004b).

The linear component of the model contains the de sign matrix and the vector of parameters to be estimated. The design matrix of independent variables, X, is composed of N rows and K+1 columns. There is one parameter corresponding to each of the K columns of independent variable settings in X, plus one β0 the intercept.

The logistic regression model equates the logit transform, the log-odds of probability of success, to the linear component:

$$\text{Log}\left(\frac{\pi_i}{1-\pi_i}\right) = \sum_{k=0}^{k} x_{ik}\,\beta_k \quad i=1, 2 \ldots N \qquad (9)$$

Parameter Estimation

The goal of logistic regression is to estimate the K + 1 unknown parameters β in Eq (9). This is done with maximum likelihood estimation which entails finding the set of parameters for which the probability of the observed data is greatest. The maximum likelihood equation is derived from the probability distribution of the dependent variable(Czepiel 2010). Since each yi represents a binomial count in the ith population, the joint probability density function of Y is:

$$F(y/\beta) = \prod_{i=1}^{N} \frac{n_i!}{y_i!(n_i-y_i)!}\,\pi_i^{y_i}(1-\pi_i)^{n_i-y_i} \qquad (10)$$

For each population, there are $\binom{n_i}{y_i}$ different ways to arrange yi successes from among ni trials. Since the probability of a success for any one of the ni trials is πi, the probability of yi successes is $\pi_i^{y_i}$ Likewise, the probability of $n_i - y_i$ failures is $(1-\pi_i)^{n_i-y_i}$

The joint probability density function in Eq. (10) expresses the values of y as a function of known, fixed values for β. (Note that β is related to π by Eq. (9). The likelihood function has the same form as the probability density function, except that the parameters of the function are reversed: the likelihood function expresses the values of β in terms of known, fixed values for y. Thus,

$$L(\beta|y) = \prod_{i=1}^{N} \frac{n_i!}{y_i!(n_i-y_i)!}\,\pi_i^{y_i}(1-\pi_i)^{n_i-y_i} \qquad (11)$$

The maximum likelihood estimates are the values for β that maximize the likelihood function in Eq. (11). The critical points of a function (maxima and minima) occur when the first derivative equals 0. If the second derivative evaluated at that point is less than zero, then the critical point is a maximum Thus, finding the maximum likelihood estimates requires computing the first and second derivatives of the likelihood function. Attempting to take the derivative of Eq. (11) with respect to β is a difficult task due to the complexity of multiplicative terms. Fortunately, the likelihood equation can be considerably simplified.

First, note that the factorial terms do not contain any of the πi As a result, they are essentially constants that can be ignored: maximizing the equation without the factorial terms will come to the same result as if they were included. Second, note that since ax-y=ax/ay and after rearranging terms, the equation to be maximized can be written as:

$$\prod_{i=1}^{N}\left(\frac{\pi_i}{1-\pi_i}\right)^{y_i}(1-\pi i)_{ni} \tag{12}$$

Note that after taking e to both sides of Eq. (9),

$$\left(\frac{\pi i}{1-\pi i}\right) = e^{\sum_{k=0}^{K} x_{ik}\beta_k} \tag{13}$$

Which, after solving for πi becomes,

$$\pi i = \left(\frac{e^{\sum_{k=0}^{K} x_{ik}\beta_k}}{1+e^{\sum_{k=0}^{K} x_{ik}\beta_k}}\right) \tag{14}$$

Substituting Eq. (10) for the first term and Eq. (11) for the second term, Eq. (12) becomes:

$$\prod_{i=1}^{N}\left(e^{e^{\sum_{k=0}^{K} x_{ik}\beta_k}}\right)_{yi}\left(1-\frac{e^{\sum_{k=0}^{K} x_{ik}\beta_k}}{1+e^{\sum_{k=0}^{K} x_{ik}\beta_k}}\right)^{n_i} \tag{15}$$

Use ax-y=ax/ay to simplify the first product and replace 1 with $\frac{e^{\sum x\beta}}{e^{\sum x\beta}}$ to simplify the second product. Eq. (15) can now be written as:

$$\prod_{i=1}^{N}\left(e^{y_i\sum_{k=0}^{K} x_{ik}\beta_k}\right)\left(1+e^{\sum_{k=0}^{K} x_{ik}\beta_k}\right)^{-n_i} \tag{16}$$

This is the kernel of the likelihood function to maximize. However, it is still cumbersome to differentiate and can be simplified a great deal further by taking its log. Since the logarithm is a monotonic function, any maximum of the likelihood function will also be a maximum of the log likelihood function and vice versa. Thus, taking the natural log of Eq. (16) yields the log likelihood function:

$$l(\beta) = \sum_{i=1}^{N} y_i\left(\sum_{k=0}^{K} x_{ik}\beta_k\right) - n_i.\log\left(1+e^{\sum_{k=0}^{K} x_{ik}\beta_k}\right) \tag{17}$$

To find the critical points of the log likelihood function, set the first derivative with respect to each β equal to zero. In differentiating Eq. (17), noting that

$$\frac{\partial}{\partial\beta_k}\sum_{k=0}^{K} x_{ik}\beta_k = x_{ik} \tag{18}$$

Since the other terms in the summation do not depend on βk and can thus be treated as constants. In differentiating the second half of Eq. (17), taking note of the general rule that $\frac{\partial}{\partial x}\log y = \frac{1}{y}\frac{\partial y}{\partial x}$ Thus, differentiating Eq. (17) with respect to each βk,

$$\frac{\partial l(\beta)}{\partial \beta_k} = \sum_{i=1}^{N} y_i\, x_{ik} - n_i * \frac{1}{1+e^{\sum_{k=0}^{K} x_{ik}\beta_k}} * \frac{\partial}{\partial \beta_k}\left(1 + e^{\sum_{k=0}^{K} x_{ik}\beta_k}\right)$$

$$= \sum_{i=1}^{N} y_i\, x_{ik} - n_i * \frac{1}{1+e^{\sum_{k=0}^{K} x_{ik}\beta_k}} * e^{\sum_{k=0}^{K} x_{ik}\beta_k} * \frac{\partial}{\partial \beta_k} \sum_{k=0}^{K} x_{ik}\,\beta_k$$

$$= \sum_{i=1}^{N} y_i\, x_{ik} - n_i * \frac{1}{1+e^{\sum_{k=0}^{K} x_{ik}\beta_k}} * e^{\sum_{k=0}^{K} x_{ik}\beta_k} * x_{ik}$$

$$= \sum_{i=1}^{N} y_i\, x_{ik} - n_i \pi_i x_{ik} \tag{19}$$

The maximum likelihood estimates for β can be found by setting each of the K + 1 equations in Eq. (19) equal to zero and solving for each βk.

Each such solution, if any exists, specifies a critical point–either a maximum or a minimum. The critical point will be a maximum if the matrix of second partial derivatives is negative definite; that is, if every element on the diagonal of the matrix is less than zero . Another useful property of this matrix is that it forms the variance covariance matrix of the parameter estimates. It is formed by differentiating each of the K + 1 equations in Eq. (19) a second time with respect to each element of β, denoted by βk'. The general form of the matrix of second partial derivatives is

$$\frac{\partial^2 l(\beta)}{\partial \beta_k \partial \beta_{k'}} = \frac{\partial}{\partial \beta_{k'}} \sum_{i=1}^{N} y_i\, x_{ik} - n_i \pi_i x_{ik}$$

$$= \sum_{i=1}^{N} -n_i x_{ik} \pi_i$$

$$= -\sum_{i=1}^{N} n_i x_{ik} \frac{\partial}{\partial \beta_{k'}} \left(\frac{e^{\sum_{k=0}^{K} x_{ik}\beta_k}}{1+e^{\sum_{k=0}^{K} x_{ik}\beta_k}}\right) \tag{20}$$

To solve Eq. (20) we will make use of two general rules for differentiation. First, a rule for differentiating exponential functions:

$$\frac{d}{dx} e^{u(x)} = e^{u(x)} * \frac{d}{dx} u(x) \tag{21}$$

In our case, let u(x) = $\sum_{k=0}^{K} x_{ik}\beta_k$. Second, the quotient rule for differentiating the quotient of two functions:

$$\left(\frac{f}{g}\right)'(a) = \frac{g(a).f'(a) - f(a).g'(a)}{[g(a)]^2} \tag{22}$$

Applying these two rules together allows us to solve Eq. (20).

$$\frac{d}{dx}\frac{e^{u(x)}}{1+e^{u(x)}} = \frac{(1+e^{u(x)}) * e^{u(x)} \frac{d}{dx} u(x) - e^{u(x)} * e^{u(x)} \frac{d}{dx} u(x)}{(1+e^{u(x)})^2}$$

$$= \frac{e^{u(x)} \frac{d}{dx} u(x)}{(1+e^{u(x)})^2}$$

$$= \frac{e^{u(x)}}{1+e^{u(x)}} * \frac{1}{1+e^{u(x)}} * \frac{d}{dx}\, u(x) \tag{23}$$

Thus, Eq. (20) can now be written as:

$$-\sum_{i=1}^{N} n_i x_{ik} \pi_i (1 - \pi_i) x_{ik'} \tag{24}$$

III. RESEARCH FINDINGS AND DISCUSSION

3.1. Distribution of household consumption:

Table 1: Participant's social demographics and economic distribution

	n	%
Over all	3042	100%
Woman's age		
15-24	443	14.6
25-34	1276	42
35-49	1323	43.5
Woman's education		
Less than secondary	2744	90.2
Secondary +	298	9.8
Woman's marital status		
Married/union	2499	82.2
Divorced/separated/widowed	543	17.8
Never married	0	0
Woman's employment		
Employed	381	12.5
Unemployed, agriculture	2658	87.4
Woman's perception of violence		
Acceptable	1744	57.3
Unacceptable	1297	42.7
Husband's education		
Less than secondary	2640	86.8
Secondary +	402	13.2
Husband's occupation		
Employed	858	28.3
Unemployed, agriculture	2176	71.7
Husband's age		
15-24	165	5.4
25-34	996	32.7
35-44	691	22.7
45+	1190	39.1
Husband's # sex partners last 12 months		
0-1	1928	65.8
2 +	1002	34.2
Husband's perception of violence		
Acceptable	408	13.4
Unacceptable	2634	86.6
Residence		
Urban	409	13.4
Rural	2633	86.6
Household wealth		
Bottom quintile	1886	62
Not bottom quintile	1155	38
Average number of people per sleeping room		
1	212	7.1
2_3	2368	77.9
4+	462	15.2
Wife's earning in relation to husband		
More than him	156	9.8
Less than him	1048	66.4
Same as him	304	19.2
Husband doesn't work	49	3
Other	25	1.6

The above table, shows the percent distribution of all covariates that have been identified as predictors of Intimate Partner Violence using the developed conceptual framework from variety of literatures. The analysis includes a total number of 5008 women of which 3042 (60.7%) have been interviewed about Intimate Partner Violence and their partners answered questionnaire.

85% of the study respondents are above 25 years of age against 94.6% of their partners, where the age distribution in the two age intervals 25-34 and 35-49 is almost the same for women with corresponding percentages 42% and 43.5% respectively. Whereas for their partners, a big number exceed 49 years old (39%).

Only 9.2% of responded women have completed secondary school i.e their education level is less secondary school whereas only 13.3% of their partners completed secondary school.

82.2% of respondents live together with their partners; this includes married women or cohabitating women.

Only 12.5 % of respondents employed against 28.3 % of their partners, the corresponding compliments including unemployed and famers.

Perception of violence among respondents corresponds to 43.7 % of women responding that violence toward them is unacceptable against 86.6% of their partners. Revelling that women are not aware of their wright compared to their partners.

3.2. Intimate Partner Violence period prevalence:

Table 2: IPV period prevalence

	N	%
Total number of women responded about IPV	3042	100
Experienced Intimate Partner Violence		
Moderate	1007	33%
Severe	1580	51.9
IPV in the last 12 months	1718	56.6

More than a half of women interviewed (56.6 percent) reported that they had been victims of either physical or sexual violence at least once during the past 12 months. Whether physical or sexual, the severity of violence was classified into two groups. Over a half of women experienced violence (51.9%), the violence reported was Severe and 33% of them reported moderate violence.

3.3. Multicollinearity screening test:

Table 3: Collinearity screening test

	w_age	w_edu	w_occ	w_acceptipv	h_edu	h_occ	h_age	h_part	h_acceptipv	hh_res	hh_wealth	hh_room_cat	earnings
w_age	1												
w_edu	-0.03	1.00											
w_occ	-0.01	0.29	1.00										
w_acceptipv	-0.02	0.20	0.11	1.00									
h_edu	-0.03	0.44	0.22	0.16	1.00								
h_occ	0.06	0.31	0.47	0.11	0.33	1.00							
h_age	0.70	0.01	0.01	-0.01	-0.03	0.05	1.00						
h_part	-0.01	0.00	0.03	0.01	0.02	0.00	0.01	1.00					
h_acceptipv	-0.06	0.10	0.06	0.09	0.10	0.06	-0.05	0.08	1.00				
hh_res	0.04	0.35	0.33	0.12	0.26	0.41	0.03	0.00	0.03	1.00			
hh_wealth	0.07	-0.28	-0.22	-0.10	-0.27	-0.32	0.07	-0.02	-0.11	-0.31	1.00		
hh_room_cat	0.14	0.05	0.06	0.04	0.07	0.07	0.13	0.03	0.01	0.05	-0.09	1.00	
earnings	-0.02	0.02	0.23	0.02	-0.02	0.03	-0.03	-0.03	0.04	0.00	-0.05	0.03	1.00

Table 3 provides the test for Multicollinearity among covariates. We are testing if there is a correlation of at least one independent variable with a combination of the other independent variables. Here the Pearson's R correlation coefficient is used. Correlations of r>0.5 are considered as collinear. The test shows us that Women's age and Husband's age are collinear; this may results in the fact that the same generation of men are more likely to marry the same generation of women. This means that Women's age and Husband's age will explain the same the outcome in the logistic model, Thus to produce parsimonious (efficient) multivariable models, and to prevent strange, unstable results, we remove one variable of collinear pairs. To remove one variable we need a judgment of which to remove. By using a chi-square test for association between Woman's age, Husband's age and experience of violence we found the p-values p-value=0.001 and p-value = 0.1359 respectively, This suggest that we drop husband's age in sake of Women's age.

3.4. Bivariate associations between social demographic characteristics and intimate partner violence in Rwanda:

Table 4: Bivariate associations between social-economic and demographic characteristics and intimate partner violence in Rwanda, 2010 DHS

	No violence (%)	Any violence (%)	p-value
Woman's age			0.0010
15-24	15.48	1.13	
25-34	43.15	4.32	
35-49	32.24	3.69	
Woman's education			0.0041
Less than secondary	82.05	8.54	
Secondary +	8.82	.59	
Woman's marital status			0.000
Married/union	90.37	9.63	
Divorced/separated/widowed	0.00	0.00	
Never married	0.00	0.00	
Woman's employment			0.0190
Employed	12.04	.83	
Unemployed, agriculture	78.72	8.26	
Woman's perception of violence			0.0007
Acceptable	51.17	5.71	
Unacceptable	39.7	3.43	
Husband's education			0.0010
Less than secondary	80.15	8.6	
Secondary +	10.69	.56	
Husband's occupation			0.0893
Employed	25.41	1.66	
Unemployed, agriculture	65.44	7.5	
Husband's # sex partners last 12 months			0.0002
0-1	85.86	8.35	
2 +	5.01	.79	
Husband's perception of violence			0.0190
Acceptable	17.5	2.41	
Unacceptable	73.37	6.73	
Residence			0.0067
Urban	12.05	.62	
Rural	78.81	8.52	
Household wealth			0.0128
Bottom quintile	52.88	6.16	
Not bottom quintile	37.99	2.98	

Average number of people per sleeping room			0.0004
1	5.26	.44	
2-3	66.6	6.06	
4+	19.09	2.56	
Wife's earning in relation to husband			0.0354
More than him	5.29	.5	
Less than him	39.24	4.65	
Same as him	12.11	.99	
Husband doesn't work	34.2	3.01	
Other			

Table 4 shows whether individual covariate is associated with the outcome (IPV). Of all the variables included in the model, some became statistically significant and other did not. Using a threshold of α=0.05, Women's age, Women's education level, Women's marital status, Women's employment status, Women's perception of violence, Husband's education level, Husbands' number of sex partners in the last 12 months prior to the survey, Husbands' perception of violence, Residence, Household wealth, Average number of person per sleeping room, and Wife's earning in relation to husband found to be associated with women's experience of Intimate Partner Violence with corresponding p-values (p-value=0.001, p-value=0.0041, p-value= 0.000, p-value= 0.0190, p-value=0.0007, p-value=0.0010, p-value=0.0002, p-value=0.0190, p-value=0.0067, p-value=0.0128, p-value=0.0354 respectively). Whereas husbands' occupation found to not be associated with IPV (p-value=0.0893).

The table above enable us to say that experience of IPV increases with women's age from 15 years aged women to 34 years aged women and from there, it starts to decrease. It is evident that women in the age group between 15 to 34 years were significantly at one time risk of facing IPV compared to the women belonged 34 years and above. There is a slight IPV prevalence decline after 34 years and above age group, which was quite expected as women of higher age group were bound to reduce violence with the passage of time by virtue of their position betters with having adult sons in the family.

IPV decreases with the increase of Women's education level, husband's number of sex partners in the last 12 months and is prevalent among unemployed or farmer women. IPV is also prevalent among those women whose perception accepts to be victims of ipv and those women whose husbands knows that ipv is unacceptable. Women's whose husbands are unemployed or farmer are more likely to experience IPV.

IPV is prevalent among women's residing in rural areas, lying in the bottom household wealth, whose families have an average number of people per sleeping room of 2 to 3 person, Women earning less than his husband or whose husband doesn't work, Thus this variable will not be considered in Multiple logistic regression model.

3.5. Multiple logistic regression analysis:

All the covariates identified by binary logistic analysis to be associated with the outcome are going to be considered in multiple logistic regression model. First we are going to run a full mode by incorporating all significant covariates.

Table 5: Multiple logistic regression full model

	Odds Ratio	Std.Err.	t	P>|t|	[95% Conf. Interval]	
Woman's age						
15-24	**1**					
25-34	1.54	0.17	3.91	0.00	[1.24	1.91]
35-49	1.53	0.21	3.14	0.00	[1.17	1.99]
Woman's education						
Secondary +	**1**					
Less than secondary	1.11	0.20	0.58	0.56	[0.78	1.57]
Woman's employment						
Employed						
Unemployed, agriculture	1.11	0.19	0.62	0.53	[0.80	1.54]
Woman's perception of violence						
Unacceptable	**1**					
Acceptable	1.28	0.12	2.65	0.01	[1.07	1.54]

Husband's education						
Secondary +	**1**					
Less than secondary	1.41	0.24	2.04	0.04	[1.01	1.98]
Husband's occupation						
Unemployed, agriculture	**1**					
Employed	0.92	0.11	-0.70	0.49	[0.73	1.16]
Husband's # sex partners last 12 months						
0-1	**1**					
2 +	2.21	0.53	3.31	0.00	[1.38	3.55]
Husband's perception of violence						
Acceptable	**1**					
Unacceptable	1.24	0.15	1.81	0.07	[0.98	1.57]
Residence						
Urban	**1**					
Rural	1.20	0.19	1.13	0.02	[1.08	1.64]
Household wealth						
Bottom quintile	**1**					
Not bottom quintile	0.88	0.10	-1.14	0.25	[0.71	1.10]
Average number of people per sleeping room						
1	**1**					
2_3	1.54	0.29	2.31	0.02	[1.07	2.22]
4+	1.86	0.38	3.03	0.00	[1.24	2.78]
Wife's earning in relation to husband						
More than him	**1**					
Less than him	0.55	0.13	-2.52	0.01	[0.34	0.88]
Same as him	0.49	0.13	-2.68	0.01	[0.29	0.83]
Husband doesn't work	0.51	0.13	-2.73	0.01	[0.31	0.83]
Constant	0.49	0.19	-1.85	0.07	[0.23	1.05]

In the multivariate logistic regression analysis, we found that for women whose age lying in 25 to 34 age interval or 35 and above (OR 1.54; 95% CI :1.24-1.91; p-value <0.01 and OR 1.53; 95% CI: 1.17-1.99; p-value<0.01 respectively), having perceptions that accepting intimate partner violence (OR 1.28; 95% CI :1.07-1.54; p-value=0.01), having a husband whose education level is less that secondary (OR 1.41; 95% CI: 1.01-1.98; p-value<0.01) or a husband who had multiple sex partners in the last 12 months prior to the survey (OR 2.21; 95% CI: 1.38-3.55, p-value<0.001), and women whose household's average number of people per sleeping room exceeds one person (2_3 persons : OR 1.54; 95% CI:1.07-2.22; p-value=0.02 and 4+ persons: OR 1.86; 95% CI: 1.24-2.78; p-value<0.01), and women living in Rural areas (OR 1.2; 95% CI: 1.08-1.64, p-value=0.02) were statistically significant risk factors for women's exposure to intimate partner violence. Whereas for wives earning less than or same as her husband or whose husband does not work (OR 0.55; 95% CI: 0.34-0.88; p-value =0.01, OR 0.49; 95% CI: 0.29-0.83; p-value =0.01, and OR 0.51; 95% CI: 0.31-0.83; p-value=0.01 respectively) was statistically significant protective factors against Intimate Partner Violence. Husband employment, Household wealth, and Husband perception of Intimate Partner Violence found to have no association with IPV (OR 0.92; 95% CI: 0.73-1.16, p-value = 0.49, OR 0.88; 95% CI: 0.71-1.10, p-value = 0.25, and OR 1.24; 95% CI: 0.98-1.57, p-value = 0.07 respectively).

Performing a manual backward stepwise logistic regression by removing any covariate which shows no significance, we finally get the following reduced model as illustrated by the following table.

Table 6: multiple logistic regression reduced model

	Odds Ratio	Std.Err.	T	P>|t|	[95% Conf.	Interval]
Woman's age						
15-24	**1**					
25-34	1.50	0.16	3.77	0.00	[1.21	1.85]
35-49	1.49	0.20	3.00	0.00	[1.15	1.94]
Woman's perception of violence						
Unacceptable	**1**					

Acceptable	1.28	0.12	2.63	0.01	[1.06	1.54]
Husband's education						
Less than secondary	**1**					
Secondary +	0.66	0.11	-2.55	0.01	[0.48	0.91]
Husband's # sex partners last 12 months						
0-1	**1**					
2 +	2.31	0.54	3.57	0.00	[1.46	3.67]
Average number of people per sleeping room						
1	**1**					
2_3	1.56	0.29	2.42	0.02	[1.09	2.24]
4+	1.86	0.38	3.06	0.00	[1.25	2.76]
Residence						
Urban	**1**					
Rural	1.24	0.18	1.50	0.04	[1.03	1.65]
Constant	0.39	0.09	-4.27	0.00	[0.25	0.60]

Table 6 provides the final model with all covariates being statistically significant and the overall significance of the model test shows that the model fits well the data and is significant with p-value < 0.0001.

Women whose age exceeds 25 years are more likely to experience intimate partner violence compared with women below 25 years old. Moreover, the risk of experiencing Intimate Partner violence increases with age from 15 years to 34 years (OR 1.5; 95% CI: 1.21-1.8; p-value<0.01) and start to decrease slightly for older women 35 years and above (OR 1.49; 95% CI: 1.15-1.94; p-value<0.01). Which was quite expected as women of higher age group were bound to reduce violence with the passage of time by virtue of their position betters with having adult sons in the family?

The risk of women to experience Intimate Partner Violence increases with knowledge of women's rights. Woman's perception of violence becomes statistically significant showing that women's IPV perceptions influence experience of IPV for women by their intimate partners. The model shows that women who knows that IPV is accepted are more likely to experience IPV compared to those who know that IPV is not accepted (OR 1.28; 95% CI: 1.06-1.54; p-value=0.01).

Husband's education level became a protective factor for women to experience IPV. Women whose partners' education level is at least secondary school are less likely to face IPV compared to those women whose husbands' education level is less than secondary school (OR 0,66; 95% CI: 0.48-0.91; p-value=0.01).

Multiple sex partners in this context means husbands who have more than on sex partner. Husband's number of sex partners in the last 12 months prior to the survey became statistically significant revealing that there is a difference in experiencing IPV between women whose husbands had multiple sex partners in the last 12 months prior to the survey compared to those women whose husband did not. The model shows that women whose husbands had had multiple sex partners in the last 12 months prior to the survey are more than two times more likely to face IPV compared to those whose husband did not (OR 2.31; 95% CI: 1.46-3.67; p-value=0.01).

Sleeping room density is associated with occurrence of IPV in the household, the risk of experience of IPV for women increases with sleeping room density. Women living in household with average number of people per sleeping room of more than one person are more likely to face IPV compared to those with one person per sleeping room on average. The model shows that women whose household's average number of people per sleeping room of 2 to 3 or 4 and above has a greater risk of experiencing IPV compared to those women whose household's average number of people per sleeping room is on person (OR 1.56; 95% CI: 1.09-2.24; p-value=0.02, OR 1.86; 95% CI: 1.25-2.76; p-value<0.01 respectively).

Residence also found to be associated with IPV. Depending on whether the woman live in Rural or urban areas influence differently on IPV experience. The modes illustrates that women living in rural areas are more exposed to IPV compared to those women living in urban areas (OR 1.24; 95% CI: 1.03-1.65; p-value=0.04).

IV. CONCLUSION AND RECOMMENDATIONS

4.1. Conclusion:

Bivariate analysis was conducted to identify among the expected predictors of IPV in Rwanda which individually contributes to the occurrence of IPV in the country. A chi-square test for association was used with the specified threshold (α=0.05).

Women's age, Women's education level, Women's marital status, Women's employment status, Women's perception of violence, Husband's education level, Husbands' number of sex partners in the last 12 months prior to the survey, Husbands' perception of violence, Residence, Household wealth, Average number of person per sleeping room, and Wife's earning in relation to husband shows a statistically significant association with women's experience of Intimate Partner Violence with corresponding p-values (p-value=0.001, p-value=0.0041, p-value= 0.000, p-value= 0.0190, p-value=0.0007, p-value=0.0010, p-value=0.0002, p-value=0.0190, p-value=0.0067, p-value=0.0128, p-value=0.0354 respectively). Whereas husbands' occupation shows no association with IPV (p-value=0.0893).

These statistics are consistent with the ones provided by the United Nations country assessment on violence against women 2013 report where, men found to be culturally trained to be breadwinners, with women playing a more subservient role. Therefore, men find it a challenge to accept women's earning capacities as this is likely to challenge their powers; thus Wife's earning in relation to husband become an IPV predictor.

All covariates showed a statistically significant association with Intimate Partner violence were included in the multiple logistic regression analysis model to see the overall contribution of factors to the outcome of interest. The final model with all covariates being statistically significant and the overall significance of the model test shows that the model fits well the data and is significant with p-value < 0.0001.

Women whose age exceeds 25 years are more likely to experience intimate partner violence compared with women below 25 years old. Moreover, the risk of experiencing Intimate Partner violence increases with age from 15 years to 34 years (OR 1.5; 95% CI: 1.21-1.8; p-value<0.01) and start to decrease slightly for older women 35 years and above (OR 1.49; 95% CI: 1.15-1.94; p-value<0.01). Which is quite consistent with a current study done in one east African country, Kenya in Kisumu district where Uwayo et al shows that young women whose ages ranges from 15 to 24 years old were more likely to experience IPV in all its forms than those with older ages.

The risk of women to experience Intimate Partner Violence increases with knowledge of women's rights. Woman's perception of violence becomes statistically significant showing that women's IPV perceptions influence experience of IPV for women by their intimate partners. The model shows that women who knows that IPV is accepted are more likely to experience IPV compared to those who know that IPV is not accepted (OR 1.28; 95% CI: 1.06-1.54; p-value=0.01). This confirms with the study conducted by the National Commission for Unity and Reconciliation (NURC) in Rwanda, where ignorance, cultural norms and practices were shown to be predictors of violence.

Husband's education level became a protective factor for women to experience IPV. Women whose partners' education level is at least secondary school are less likely to face IPV compared to those women whose husbands' education level is less than secondary school (OR 0,66; 95% CI: 0.48-0.91; p-value=0.01), which was shown by a secondary analysis of KDH survey where violence was significantly lower for women whose partners had attained at least a postsecondary education(Garcia-Moreno et al. 2012).

Multiple sex partners in this context means husbands who have more than on sex partner. Husband's number of sex partners in the last 12 months prior to the survey became statistically significant revealing that there is a difference in experiencing IPV between women whose husbands had multiple sex partners in the last 12 months prior to the survey compared to those women whose husband did not. The model shows that women whose husbands had had multiple sex partners in the last 12 months prior to the survey are more than two times more likely to face IPV compared to those whose husband did not (OR 2.31; 95% CI: 1.46-3.67; p-value=0.01). This confirms with Kenyan study where Diane et all reported that women indicated that their male partner's multiple sexual partners habit was a factor which contributed to their abuse. This sentiment was espoused in the women's focus group in Obunga where one woman noted "some men leave work with little money and instead of bringing it [home] they go and pay for sex with sex workers or in other kinds of leisure with their concurrent partners(Uwayo 2014).

Sleeping room density is associated with occurrence of IPV in the household, the risk of experience of IPV for women increases with sleeping room density. Women living in household with average number of people per sleeping room of more than one person are more likely to face IPV compared to those with one person per sleeping room on average. The model shows that women whose household's average number of people per sleeping room of 2 to 3 or 4 and above has a greater risk of experiencing IPV compared to those women whose household's average number of people per sleeping room is on person (OR 1.56; 95% CI: 1.09-2.24; p-value=0.02, OR 1.86; 95% CI: 1.25-2.76; p-value<0.01 respectively).

Residence also found to be associated with IPV. Depending on whether the woman live in Rural or urban areas influence differently on IPV experience. The modes illustrates that women living in rural areas are more exposed to IPV compared to those women living in urban areas (OR 1.24; 95% CI: 1.03-1.65; p-value=0.04). This shows the same results as UN country assessment 2013 report on Violence Against Women (VAW) that reported that rural women are recognized as being at special risk of violence due to the prevalence of traditional attitudes in many rural communities(United Nations & UNFPA 2013).

4.2. Recommendations and Suggestions:

1. Based on my findings and previous research on intimate partner violence, IPV is still a major problem in Rwanda. I recommend the following:
2. Continue efforts in raising awareness among and training of both men and women around human rights with a specific focus on women's rights at all levels in the communities.
3. Recognize, expand and increase access to services that improve women's decision-making ability in response to IPV, particularly counselling and support groups that provide benefits through discussion with others.
4. Expand opportunities for collective discussions about broader gender issues and attitudes that contribute to IPV.
5. At the individual level, continue to support women in seeking assistance from the police, while linking this more strongly to additional supports for those in both married and unmarried relationships, and for communities.
6. As for future research recommendations, due to time constraints this study did not cover all predictors of IPV in Rwandan context, it only used RDHS 2010 data set and limited on those factors collected by the survey. Thus further studied would explore deeply all other factor that did not covered by the current study.

REFERENCES

[1] Abramsky, T. et al., 2011. What factors are associated with recent intimate partner violence? findings from the WHO multi-country study on women's health and domestic violence. BMC public health.

[2] Balian, S. et al., 2014. Understanding and Using the Demographic and Health Surveys DHS Curriculum Facilitator's Guide About the DHS Curriculum Facilitator's Guide.

[3] Bazargan-Hejazi, S. et al., 2013. Patterns of intimate partner violence: a study of female victims in Malawi. Journal of injury & violence research.

[4] Boe, P., 2004. DHS Comparative Reports,

[5] Bonomi, A.E. et al., 2006. Intimate Partner Violence and Women's Physical, Mental, and Social Functioning. American Journal of Preventive Medicine.

[6] Czepiel, S., 2010. Maximum likelihood estimation of logistic regression models: theory and implementation. Avaiable at http://czep.net/stat/mlelr.pdf,

[7] DHS, 2011. Demographic and Health Survey 2010 Rwanda,

[8] DHS, 2006. Rwanda Demographic and Health Survey, Available at: http://www.measuredhs.com.

[9] Dr.Avishek Mallick, 2014. Logistic regression — modeling the probability of success. In Modeling the probability of success. p. 71.

[10] Garcia-Moreno, C. et al., 2012. Prevalence of intimate partner violence: findings from the WHO multi-country study on women's health and domestic violence. Lancet, 368(9543), pp.1260–9. Available at: http://www.thelancet.com/ article/S0140673606695238/fulltext [Accessed February 20, 2015].

[11] Gillard, J.W., 2012. Method of moments estimation in linear regression with errors in both variables.

[12] GMO, 2013. Gender Monitoring Office Annual Report 2012-2013,

[13] GMO, 2005. REPUBLIC OF RWANDA Gender Monitoring Office (GMO) GENDER BEST PRACTICES IN RWANDA,

[14] Hellevik, O., 2009. Linear versus logistic regression when the dependent variable is a dichotomy. Quality and Quantity.

[15] Hindin, M.J., 2008. DHS ANALYTICAL STUDIES 18.

[16] Jewkes, R. et al., 2006. Factors associated with HIV sero-status in young rural South African women: Connections between intimate partner violence and HIV. International Journal of Epidemiology.

[17] Koenig, M.A. et al., 2003. Domestic violence in rural Uganda: Evidence from a community-based study. Bulletin of the World Health Organization.

[18] Mannell, J. & Jackson, S., 2014. Intimate Partner Violence in Rwanda: Women's voices.

[19] Ntaganira, J. et al., 2008. Intimate partner violence among pregnant women in Rwanda. BMC women's health.

[20] Pohar, M., Blas, M. & Turk, S., 2004a. Comparison of Logistic Regression and Linear Discriminant Analysis: A Simulation Study. Metodološki zvezki.

[21] Pohar, M., Blas, M. & Turk, S., 2004b. Comparison of Logistic Regression and Linear Discriminant Analysis: A Simulation Study. Metodološki zvezki, 1(1), pp.143–161. Available at: http://mrvar.fdv.uni-lj.si/pub/mz/mz1.1/pohar.pdf.

[22] Project, M.A.P., 2011. Modeling Conditional Probabilities 1 : Lucky Dip Modeling Conditional Probabilities 1 : Lucky Dip. In Lucky Dip. p. 17.

[23] Sampath, V., Flagel, A. & Figueroa, C., 2011. A LOGISTIC REGRESSION MODEL TO PREDICT FRESHMEN ENROLLMENTS. In A LOGISTIC REGRESSION MODEL TO PREDICT FRESHMEN ENROLLMENTS. pp. 36–42.

[24] Turner, H., 2008. Introduction to Generalized Linear Models.

[25] United Nations & UNFPA, 2013. Country Assessment on Violence against Women Rwanda,

[26] Uwayo, D., 2014. Factors Contributing to Intimate Partner Violence and the Effectiveness of Services Available to Help Victims in Kisumu, Kenya .

[27] WHO, 2012. Understanding and addressing violence against women Sexual violence. , p.12.

[28] WHO, 2013. Understanding and Preventing Violence,