



Spatial and Multilevel Analysis on the Determinants of Gender-Based Violence Against Women in Ethiopia

Kenenisa Abdisa Kuse¹

Bule Hora University, College of Natural and Computational Science, Department of Statistics, Bule Hora Ethiopia.

Email: abdisakenenisa40@gmail.com/kenenisa.abdisa@bhu.edu.et

ABSTRACT

Violence against women and girls is a global human rights violation and a substantial development challenge. It affects women in the world and crosses cultural and economic boundaries. This research was designed to explore the distribution and variation of gender-based violence against women aged between 15 to 49 years and to identify the associated determinant factors using the evidence from the 2016 EDHS. In this study, a sample of 5860 women of reproductive age (15-49) years were selected using a two-stage stratified random sampling. The multilevel random coefficient model has a best-fitted model and estimates the variation of women's violence across the regions and zones of Ethiopia. Further, Spatial analysis was performed to analyze the spatial distribution of women's violence across the regions and zones of Ethiopia. Accordingly, At the national level, 23% of women have experienced physical violence, 19% have experienced emotional violence, and 10% have experienced sexual violence. The multilevel random coefficient model identified that there is significant evidence of variation among regions and zone clusters in Ethiopia. Similarly, spatial analysis indicated that the pattern of spatial distribution of women was randomly clustered at regional and zonal levels in Ethiopia. In general, the study indicated that women's violence is still a public health problem in Ethiopia. Thus, the government should ensure the legal framework and policies including enforcement of legal punishment for the perpetrators of gender-based violence against women.

Keywords: Physical Violence, Sexual Violence, Spatial Analysis, Multilevel Analysis, Ethiopia

1. INTRODUCTION

According to the United Nations, gender-based violence encompasses any action that leads to or is likely to cause, physical, sexual, or psychological harm or suffering to women. This includes threats, coercion, and arbitrary deprivation of liberty, whether occurring in public or private contexts” (WHO, 2021). Gender-based violence manifests in diverse forms and can be categorized based on different criteria. It may be delineated by the relationship between the perpetrator and the victim, distinguishing between intimate partner violence (IPV) and non-IPV scenarios. Alternatively, it can be classified by the type of violent act, encompassing sexual, physical, or emotional violence (World Bank, 2019).

Violence is often the cycle of abuse that manifests itself in many forms throughout women’s lives. Violence against women may include enforced malnutrition lack of access to medical care, female genital mutilation, early marriage, forced prostitution... etc. Some go on to suffer throughout their adult lives - battery, rape, and even murder at the hands of an intimate partner (Izugbara et al., 2020). Violence against women is a global human rights violation and a substantial development challenge. It affects women throughout the world and crosses cultural and economic boundaries. World Health Organization estimates that more than 30% of women worldwide have experienced either physical or sexual partner violence (Devries et al, 2013; Stöckl et al, 2013). 7% of women worldwide have experienced non-partner sexual assault (Abrahams et al., 2014). About 100–140 million women worldwide have undergone female genital mutilation (FGM) and more than 3 million are at risk for FGM every year in Africa alone (Feldman-Jacobs and Clifton 2014). Nearly 70 million worldwide have been married before the age of 18 years, many of them against their will (Degue et al., 2014; Modi et al., 2014). The effect of violence against women on their health and welfare, their families, and communities is substantial (Devries et al., 2013). The costs of violence against women, both direct and indirect, are a staggering burden for households and economies (World Bank, 2014).

According to studies in different countries on women’s health and domestic violence against women, the lifetime prevalence of physical, sexual, or both physical and sexual violence ranges from 15% (Japan) to 71% (Ethiopia). Nearly one-half (49%) of ever-married women faced physical violence, 59% of them experienced sexual violence, and 71% of them had one or the other form of

violence, or both, over their lifetime. About 35% of all ever-married women experienced at least one severe form of violence by a partner (Chernet and Cherie, 2020). A cross-sectional study conducted in Nigeria showed that almost one in four (21.5%) ever-married women faced IPV at some point in their lives (Benebo FO, 2018). According to research conducted in Gondar referral Hospital, the overall prevalence of domestic violence among pregnant women was estimated to be 58.7% with emotional violence being the most common (57.8%), followed by physical violence (32.2%), and sexual violence (7.6%). It also showed that housewives, women with no salary of their own, partners' daily use of alcohol, and women who disobeyed their partners were found to be positively and significantly associated with domestic violence during pregnancy (Fekadu, 2018).

In Ethiopia, violence against women continues to be a major challenge and a threat to women's empowerment. Women face physical, emotional, and sexual abuses that undermine their health and ability to earn a living; disrupt their social systems and relationships; and rob them of their childhood and education (CSA, 2016). Therefore, the main objective of this study was to explore the distribution and variation of gender-based violence against women aged between 15 to 49 years and to identify the associated determinant factors using the evidence from the 2016 EDHS.

2. METHODOLOGY

2.1. Data Sources

This study used data collected from the Ethiopian Demographic and Health Survey (EDHS) which was conducted by the Central Statistical Agency (CSA) under the auspices of the support of the Ministry of Health. The sampling frame used for the EDHS was the Population and Housing Census conducted by the Central Statistical Agency (CSA) in 2007. During the 2007 Population and Housing Census, each of the kebele was subdivided into convenient areas called census enumeration areas (EAs).

2.2. Sampling techniques

The EDHS sample was selected using a stratified, two-stage cluster design and EAs were the sampling units for the first stage. The interviewer-administered questionnaire was used to collect data on women of reproductive age (15-49) years. The questionnaire included socio-demographic, socioeconomic, pregnancy, and maternal health service-related variables related to women's health. A stratified two-stage cluster sampling with a total of 645 Enumeration Areas (EAs) (202 in urban

and 443 in rural areas) was selected with a probability proportional to EA size. A total of 5,860 women were asked questions about violence against women. Three percent of women eligible for the domestic violence module could not be successfully interviewed, mainly due to lack of privacy. Specially constructed weights were used to adjust for the selection of only one woman per household and to ensure that the domestic violence subsample was nationally representative.

2.3. Variables of the study

Dependent variables

The dependent variable for this study is gender-based violence. The three measures of gender-based violence against women are:

Physical Violence: pushing her, shaking her, or throwing something at her; slapping you; twisting her arm or pulling her hair; punching her with his/her fist or with something that could hurt her; kicking her, dragging her, or beating her up; trying to choke her or burning her on purpose; or threatening or attacking her with a knife, gun, or any other weapon

Sexual Violence: physically forcing a female to have sexual intercourse with him even when she did not want to; physically forcing her to perform any other sexual acts without her interest; forcing her with threats or in any other way to perform sexual acts she did not want to.

Emotional Violence: acting something to humiliate female in front of others; threatening to hurt or harm her or someone close to her; insulting or making her feel bad about herself.

Therefore, gender-based violence was captured as a dichotomous variable and coded as 1 if the female experienced any type of violence (Physical violence, sexual violence, or emotional violence), and 0 otherwise.

Independent Variables

The independent variables of the study were classified as demographic and socioeconomic variables which are expected to determine gender-based violence. The predictor variables included in the study were: Place of residence, region, employment status, educational level, age, marital status, number of living children, wealth index, husband alcohol consumption, and religion.

3.4. Methods of Data Analysis

Multilevel Logistic Regression Model

A multilevel statistical approach was used to model the relationship between gender-based violence against women status and the explanatory variables. In a multilevel logistic regression model, two levels of data hierarchy were stated (for instance, individual women and region). Units at one level are nested within units at the next higher level. In this study, the basic data structure of the two-level logistic regression is a collection of J groups (regions) and within-group j ($j= 1,2,\dots, J$), a random sample n_j of level-one units (females). The response variable is denoted by;

$$Y_{ij} = \begin{cases} \mathbf{1} & \text{if the } i^{th} \text{ female in the } j^{th} \text{ region is experienced violence} \\ \mathbf{0} & \text{if the } i^{th} \text{ female in the } j^{th} \text{ region is not experienced violence} \end{cases}$$

With probabilities, $P_{ij} = P(Y_{ij} = 1 | X_{ij}, u_{ij})$ is the probability of experiencing violence for the i^{th} youth in the j^{th} region, and $1 - P_{ij} = P(Y_{ij} = 0 | X_{ij})$ the probability of not experiencing violence for the i^{th} youth in the j^{th} region.

The Random Intercept Model

The Random intercept model is used to model unobserved heterogeneity in the overall response by introducing random effects. In the random intercept model, the intercept is the only random effect meaning that the groups differ for the average value of the response variable, but the relationship between explanatory and response variables cannot differ between groups.

The random intercept model expresses the log-odds, i.e. the logit of P_{ij} , as a sum of a linear function of the explanatory variables. That is,

$$\log\left(\frac{P_{ij}}{1 - P_{ij}}\right) = \beta_{0j} + \beta_1 x_{1ij} + \beta_{2ij} x_{2ij} + \dots + \beta_k x_{kij}, i = 1, 2, \dots, n, j = 1, 2, \dots, J$$

Where the intercept term β_{0j} is assumed to vary randomly and is given by the sum of an average intercept β_0 and group-dependent random errors U_{0j} , that is $\beta_{0j} = \beta_0 + U_{0j}$.

As a result, we have $\log it(P_{ij}) = \beta_0 + \sum_{h=1}^k \beta_h x_{hij} + U_{0j}$ where $\beta_0 + \sum_{h=1}^k \beta_h x_{hij}$ the fixed part of the model and the remaining U_{0j} is called the random part of the model. It is assumed that the residual U_{0j} is mutually independent and normally distributed with mean zero and variance δ_0^2

The Random Coefficient Model

The random coefficients build up on the random intercept model by allowing the effects of individual predictors to vary randomly across level 2, that is, level 1 slope coefficients are allowed to take on different values in different aggregate groupings. In the random coefficient model, both the intercepts and slopes are allowed to differ across the region. It is given by:

$$\log(P_{ij}) = \log\left(\frac{P_{ij}}{1 - P_{ij}}\right) = \beta_0 + \sum_{h=1}^k \beta_h x_{hij} + U_{0j} + \sum_{h=1}^k U_{1j} X_{1ij}$$

Intra-Class Correlation Coefficient (ICC): The researcher purposively used ICC to measure the reliability of ratings for the clusters. The logistic distribution for the level of one residual e_{ij} implies a variance of $\frac{\pi^2}{3} = 3.29$. Thus, in this study, the researcher studies the variation of women across Regions and Zones. The ICC for the three-level binary data can be defined for each level separately;

$$ICC(Region) = \frac{\delta^2_{Region}}{\delta^2_{Region} + \delta^2_{Zone} + \frac{\pi^2}{3}} \dots \dots \dots ICC \text{ attributable to level 3}$$

$$ICC(Zone) = \frac{\delta^2_{Zone}}{\delta^2_{Region} + \delta^2_{Zone} + \frac{\pi^2}{3}} \dots \dots \dots ICC \text{ attributable to level 2}$$

$$ICC(Zone, Region) = \frac{\delta^2_{Zone} + \delta^2_{Region}}{\delta^2_{Region} + \delta^2_{Zone} + \frac{\pi^2}{3}} \dots \dots ICC \text{ attributable to levels 2 and 3}$$

Where $\frac{\pi^2}{3}$ denotes the variation of lower (individual) level units, δ^2_{Region} and δ^2_{Zone} denote the variation between Regions and Zones respectively.

Proportional Change in Variance (PCV): It was also computed for each model for the empty model to show the power of the factors in the models in explaining the outcome variables and obtained as;

$$PCV = \frac{V_e - V_i}{V_e}$$

Whereas V_e is the variance of women's level in the empty model and V_i is variance in successive models.

Spatial Analysis

Spatial analysis is an analysis that includes the influence of spatial or space into the analysis and includes any of the formal techniques that study entities using their topological, geometric, or geographic properties (Don et al, 2008). Therefore, for this study, the researcher used spatial analysis to show the spatial distribution of gender-based violence against women across regions and zones in Ethiopia.

Spatial scan statistical analysis

It tests the presence of statistically significant spatial clusters of physical violence, sexual violence, and emotional violence using Kuldorff's SaTScan version 9.6 software. Women who had experienced physical violence, sexual violence, and emotional violence were considered cases while those who had not experienced violence were taken as controls to fit the Bernoulli model. (Coleman et al., 2009).

Spatial Model

Spatial dependence model

Spatial dependence refers to the degree of spatial autocorrelation between independently measured values observed in geographical space. The need for spatial regression models arises when the outcome of interest is correlated with the outcomes of its neighbor (conditional on other variables). If there is no spatial dependency, then we should use ordinary least squares (OLS) instead (Elhost, 2010). The two common types of spatial dependence were spatial lag models and spatial error models.

Spatial Error Model

The error terms across different spatial units are correlated. With spatial error in OLS regression, the assumption of uncorrelated error terms is violated. As a result, the estimates are inefficient. The spatial error model accounts for spatial dependency by an error term and a associated spatially lagged error term (Anselin & Bera, 1998). A spatial error model is specified as;

$$Y = XB + u, u = \rho Wu + \varepsilon$$

Whereas Y is an n-by-1 vector of response variables, X is an n-by-p design matrix of explanatory variables, B is a p-by-1 vector of regression coefficients, μ is an n-by-1 vector of error terms, ρ is a scalar spatial error parameter, W is an n-by-n spatial weight matrix, and ε is an n-by-1 vector of error terms that are normally and independently, but not necessarily identically distributed.

Spatial Lag Model

Refers to the dependent variable y in place i is affected by the independent variables in both places i and j . With spatial lag in OLS regression, the assumption of uncorrelated error terms is violated; in addition, the assumption of independent observations is also violated (Lesage and Pace, 2009). A spatial Lag model is specified as:

$$Y = XB + \rho Wy + \varepsilon$$

Whereas Y is an n-by-1 vector of response variables, X is an n-by-p design matrix of explanatory variables, B is a p-by-1 vector of regression coefficients, μ is an n-by-1 vector of error terms, ρ is a scalar spatial lag parameter, W is an n-by-n spatial weight matrix, and ε is an n-by-1 vector of error terms that are normally and independently, but not necessarily identically distributed.

Spatial Analysis Goodness of fit test

In this study, the researchers make the comparison between the spatial lag and spatial error model, and OLS using the respective log-likelihood of the maximum likelihood estimation (Catani et al., 2017). The model with the lowest AIC is the best.

There are so many tests performed to assess the spatial dependence of the model. However, in this study, the researcher used the following test of spatial diagnosis. The statistics are the simple LM test for a missing spatially lagged dependent variable (Lagrange Multiplier (lag)), the simple LM test for error dependence (Lagrange Multiplier (error)), variants of these robust to the presence of the other (Robust LM (lag) and Robust LM (error)-which tests for error dependence in the possible presence of a missing lagged dependent variable.

These four-test statistics are computed as follows;

$$LM_{error} = \frac{(e^T We/S)^2}{T} \sim X^2(1)$$

$$LM_{lag} = \frac{(e^T Wy/(e^T e/N))^2}{R} \sim X^2(1)$$

$$Robust\ LM_{error} = \frac{(e^T Wy/s - TR^{-1} e^T We/s^2)^2}{T - T^2 R^{-1}} \sim X^2(1)$$

$$Robust\ LM_{lag} = \frac{(e^T Wy/s^2 - e^T We/s^2)^2}{R-T} \sim X^2(1)$$

Whereas, $s^2 = \frac{e^T e}{N}$, $T = tr(W^2 + W^T W)$, and $R = (WX\hat{B})^T M(WX\hat{B})(e^T e/N) + tr(W^2 + W^T W)$

Whereas, \hat{B} is the OLS estimation of model parameters in the original hypothesis

Parameter Estimation

The most common method of estimating the spatial lag and spatial error model was the maximum likelihood estimation.

Maximum Likelihood Estimation of Spatial Lag Model

The underlying assumption is the normal distribution of model errors, i.e., $\varepsilon \sim N(0, \sigma^2)$.

The log-likelihood function of the spatial lag model was;

$$\ln L = \frac{\ln|I - pW| - \frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - (y - pWy - XB)'(y - pWy - XB)}{2\sigma^2}$$

The logarithm is the determinate of the $(n \times n)$ asymmetric matrix $(I - pW)$ does not tend to zero, it constrains the parameter values to their feasible range between the inverse of smallest and

largest eigenvalues of W , since for positive autocorrelation, as $p \rightarrow 1, \ln|I - pW| \rightarrow -\infty$. The first condition for the maximum likelihood estimators yields non-linear (in parameters) equations which are solved by the numerical methods (Newton Raphson iteration method). The maximum likelihood estimate of p is obtained from a numerical optimization of the concerned log-likelihood function.

$$L_{lag}^c = \frac{-n}{2} \ln \left[\frac{(e_{ols} - p_{el})'(e_{ols} - p_{el})}{n} \right] + \sum_{i=1}^n \ln(I - p\omega_i)$$

Where e_{ols} and e_l are, respectively, the residuals from OLS regressions of Y on X and WY on X and ω_i 's are the eigenvalues of the spatial weights matrix W. Given the maximum likelihood estimate of ρ , the parameters β , and the error variance σ^2 are then easily computed.

Maximum Likelihood Estimation of Spatial Error Model

The log-likelihood function of the spatial error model was;

$$\ln L = \frac{\ln|I - pW| - \frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - (y - XB)'(I - pW)'(I - pW)(y - XB)}{2\sigma^2}$$

The maximum likelihood estimation for the spatial error model employs the error term into log-likelihood function as follows:

$$L_{error} = \sum_{i=1}^n \ln|I - p\omega_i| - \frac{n}{2} \ln \left[\frac{\sigma^2 - (y - X\beta)'(I - pW)(y - X\beta)}{2\sigma^2} \right]$$

This function is then maximized over the parameter spaces. First \hat{p} is obtained and then β and σ^2 are estimated given the value of \hat{p} .

Spatial Autocorrelation

The tests are used to show the amount of spatial distribution at all. Two common measures of spatial autocorrelations are Moran's I and Gear's C tests. This paper uses Moran's I to test whether there is spatial correlation in the regions and Zones or not. Moran's I is produced by standardizing the spatial auto-covariance by the variance of the data.

Input

The input data file should contain the X, and Y coordinates and the value at each point x_i . Input whether you have a spatial weights matrix file. If do not have a spatial weights matrix, you researcher is required to enter the A and m parameters or a parameter representing the friction of distance selected. The researcher should have to enter the maximum distance, the number of steps, and whether you want bands or increments.

Analysis

The expected value of Moran's I is $-1/(N-1)$. Values of I that exceed $-1/(N-1)$ indicate positive spatial autocorrelation, in which similar values, either high values or low values are spatially clustered. Values of I below $-1/(N-1)$ indicate negative spatial autocorrelation, in which neighboring values are dissimilar. The theoretically expected value for Geary's c is 1. A value of Geary's c less than 1 indicates positive spatial autocorrelation, while a value larger than 1 points to negative spatial autocorrelation.

Formula

$$I = \frac{N \sum_{i=1}^N \sum_{j=1}^N w_{(i,j)} (x_i - \bar{x})(x_j - \bar{x})}{S_0 \sum_{i=1}^N (x_i - \bar{x})^2}, j \neq i$$

$$c = \frac{(N - 1) \sum_{i=1}^N \sum_{j=1}^N w_{(i,j)} (x_i - \bar{x})^2}{2S_0 \sum_{i=1}^N (x_i - \bar{x})^2}, j \neq i$$

Where N is the imputed value, \bar{x} is the mean of x_i , $S_0 = \sum_{i=1}^N \sum_{i=1}^N w_{(i,j)}$, $w_{(i,j)}$ is the specified weighting scheme chosen. The variances of I and c will differ according to the data model employed. Point Pattern Analysis (PPA) uses a randomization assumption. Under a randomization assumption, the variances of I and c are shown below.

$$\begin{aligned} Var(I) &= \frac{N((S_1(N^2 - 3N + 3) - (NS_2 + 3S_0^2))}{(N - 1)(N - 2)(N - 3)S_0^2} - \frac{K(S_1(N^2 - N) - 2NS_2 + 6S_0^2)}{(N - 1)(N - 2)(N - 3)S_0^2} - \left(\frac{1}{N - 1}\right)^2 \\ Var(c) &= \frac{(N - 1)S_1(N^2 - 3N + 3 - K(N - 1))}{S_0^2 N(N - 2)(N - 3)} + \frac{(N^2 - 3 - K(N - 1))^2}{N(N - 2)(N - 3)} \\ &\quad - \frac{(N - 1)S_2(N^2 + 3N - 6 - K(N^2 - N + 2))}{4N(N - 2)(N - 3)S_0^2} \end{aligned}$$

Where I indicate Moran's and C indicate Gear's test value and S_1 , S_2 , and K is expressed as follows;

$$S_1 = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (w(i, j) + w(j, i))^2$$

$$S_2 = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (w(i, j) + w(j, i))^2$$

Where the two summations indicate that the total of values in the input N by N matrix taken from the spatial weights matrix and

$$K = \frac{N \sum_{i=1}^N (x_i - \bar{x})^4}{(\sum_{i=1}^N (x_i - \bar{x}))^2}$$

Ethical Approval

The authors of this manuscript sought permission from the EDHS Program for the use of the dataset for this study. Further information about the EDHS data usage and ethical standards is available at <http://goo.gl/ny8T6X>.

3. RESULTS AND DISCUSSION

3.1. Descriptive Statistics Result

The women aged 15 to 49 years were included in the study. Accordingly, out of 5860 total, about 23% of females have experienced physical violence, 19% have experienced emotional violence, and 10% have experienced sexual violence during pregnancy at the time of the survey (Table 1). Further, cultural gender-based violence accounts for 32%. According to the respondents, the culture did not allow women to make decisions concerning their family, and 54% of the respondents agreed that FGM was a form of violence against women in their culture while 46% disagreed with the notion that FGM was a form of violence against women in their culture. 28% of the respondents agreed that it was culturally right for a husband to beat a wife. This result is consistent with the study done by. This result is consistent with the study done by (Cynthia et al., 2020 and Jepkoech, 2021).

Table 1: Descriptive Result

Gender-based Violence status	Yes	No
Physical Violence	1348 (23%)	4512 (77%)
Sexual violence	586 (10%)	5274 (90%)
Emotional violence	1114 (19%)	4746 (81%)
Culturally allow women to make decisions	3984 (68%)	1876 (32%)
FGM is a form of violence against women	3164 (54%)	2696 (46%)
culturally right for a husband to beat a woman	1641 (28%)	4219 (72%)

Source: EDHS, 2016

Multilevel Regression Results

Chi-square test

A two-level structure (with individual females as a first-level unit, and region as a second-level unit) was used. The chi-square test was used to assess heterogeneity between the regions of Ethiopia. The test results are $\chi^2 = 526$ with d.f = 10 ($p = 0.001$). This shows there is confirmation of heterogeneity to the regions.

Multicollinearity test

Before making any statistical inference, the researcher needs to check that the model fits sufficiently well check for influential observations that have an impact on the estimates of the coefficients, and make a valid analysis as follows:

Table 2: Multicollinearity checking between each explanatory variable

Variables	VIF
Place of residence	2.545
Age	2.457
Religion	1.235
Employment status	1.878
Marital status	3.335
Number of living children	4.120
husband’s alcohol consumption	2.354
Wealthy index	1.854
Region	2.142

Source: EDHS, 2016

The multicollinearity among each explanatory variable was tested using the Variance Inflation Factor (VIF). In this study, the VIF for each of the explanatory variables was less than five (5), showing the absence of multicollinearity in the models i.e. indicates that there is no problem of multicollinearity in the data.

Variance Component Model

The variance components model (i.e., a simple model without predictors) that forecasted the likelihood of violence status fitted. The fixed effect was disclosed by the variance components model results (Table 3). The estimated average log odds of violence status among females between the ages of 15 to 49 is -0.564. The figure showed a significant value and indicated the overall percentage of gender-based violence prevalence among Ethiopian women between the ages of 15 to 49.

The intraregional correlation coefficient (ICC) helps to understand the impact of regional factors on gender-based violence against women aged 15 to 49. According to a variance components model, the ICC is calculated as 0.115. The result indicates that approximately 11.5% of the overall variability in gender-based violence might be attributed to regional factors. The remaining 88.5% of the variation was associated with individual-level differences within regions. In general, regional variation shows a significant effect, but individual characteristics/factors/ also contribute significantly to gender-based violence.

Table 3: Estimates for variance components model

Fixed part	Estimate	S. error	z-value	p-value
intercept	-0.564	0.121	-4.661	0.0023
Random effect	Estimate	S. error	z-value	p-value
$\hat{\sigma}_u^2$	0.283	0.105	2.74	0.025
ICC (ρ)	0.115	.0478	2.071	

Source: EDHS, 2016

Random Intercept Model

A multilevel binary logistic model with random intercept and fixed explanatory variables was estimated to identify the effect of explanatory variables. Accordingly, the estimate between-region variance decreased from the variance component model 0.283 to the random intercept model 0.235. the result, suggesting that the distribution of fixed explanatory variables is somewhat different across

regions of the country. The results from the random intercept model (Table 4) showed that the random intercept is significant implying that the average proportion of gender-based violence against women aged between 15 and 49 years differs from region to region.

Table 4: Estimates of random intercept model

Variables	Categories	β	S.E	Wald	sign	odds
Place of residence	Urban (Ref)	----	----	----	----	----
	Rural	0.536	0.129	4.155	0.000*	3.56
Age	15-24 (Ref)	----	----	----	----	----
	25-34	0.198	0.098	2.21	0.037*	1.130
	35-49	-0.382	0.152	-2.513	0.031*	0.843
Educational Level	No education (Ref)	----	----	----	----	----
	Primary	-0.299	0.075	-3.97	0.000*	0.865
	Secondary and above	-0.642	0.109	-0.589	0.002*	0.624
Religion	Orthodox (Ref)	----	----	----	----	----
	Catholic	-0.45	0.32	-1.41	0.113	0.812
	Protestant	-0.35	0.23	-1.52	0.085	0.685
	Muslim	-0.61	0.497	-1.22	0.093	0.521
	Traditional	0.872	0.230	3.79	0.00*	1.97
	Other	0.574	0.328	1.75	0.072	1.23
Employment status	Employed (Ref)	----	----	----	----	----
	Not employed	0.924	0.27	3.42	0.032*	2.53
Marital status	Single (Ref)	----	----	----	----	----
	Married	0.499	0.116	4.30	0.025*	2.071
	Separated	0.689	0.13	5.30	0.012*	3.252
Number of lining children	0 (Ref)	----	----	----	----	----
	1-2	0.125	0.031	4.1	0.004*	2.52
	3-4	0.82	0.25	3.25	0.015*	1.64
	5+	0.93	0.314	2.96	0.037*	1.19
Wealthy index	Poor (Ref)	----	----	----	----	----
	Medium	-0.604	0.12	-5.33	0.004*	0.724
	Rich	-0.31	0.092	-3.37	0.008*	0.652
Husband's alcohol consumption	Does not drink	----	----	----	----	----
	Drink/never get drunk	0.251	0.071	3.54	0.014*	1.256
	Gets drunk sometimes	0.682	0.125	5.456	0.001*	3.64
	Gets drunk very often	0.793	0.131	6.053	0.001*	5.19
Random effect						
	$\hat{\sigma}_u^2$ (Region)	0.235	0.095			
	$\hat{\sigma}_u^2$ (Zone)	0.569	0.0012			
	ICC(Region)	0.057				
	ICC(Zone)	0.138				

Source: EDHS, 2016

The intraregional correlation coefficient (ICC) is estimated as $\hat{\rho}=0.099$. The result showed that 9.9% of the total variability in gender-based violence against women aged between 15 to 49 years is attributable to the regional level.

The result of random intercept reveals that females who live in rural areas of the country were about 35.2% more likely to experience violence (OR = 3.52) compared to females who live in urban areas, controlling for other variables in the model. The Female in the age group 25-34 years were 1.13 (OR= 1.130) times more likely to experience violence compared to females in the age group 15-24 years. However, the Female in the age group 35-49 years were about 15.7% (OR= 0.843) less likely to experience violence compared to females in the age group 15-24 years. Similarly, females following traditional religious views were about 19.7% (OR= 1.97) more likely to experience violence compared to females with orthodox religions. Employed females were about 45.6% (OR= 0.54) less likely to experience violence compared to females who were not employed.

The odds of gender-based violence against women aged between 15 to 49 years with primary and secondary education were about 13.5% and 37.6% less likely to experience violence compared to females who have no education, respectively. This finding is consistent with the previous studies done by (Ömer & ŞeydaÜnver, 2021). Regarding marital status, the women who have married and divorced were about 20.71% and 32.52% more likely to experience violence compared to a single woman, respectively. This result maintains what is stated in studies done by (Rasha & Rania, 2017). Regarding the number of children they have, women who have 1-2 children, 3-4 children, and 5+ children were about 25.2%, 16.4%, and 11.9% more likely to experience violence compared to women who have no living children, respectively. This finding is consistent with the previous studies done by (Jean et al., 2019). The incidence of gender-based violence among rich and medium females decreased by 34.8% and 27.6. Similarly, a woman whose husband can drink/never get drunk gets drunk sometimes, and gets drunk very often were about 12.56%, 36.4%, and 51.9% more likely to experience violence compared to a woman whose husband does not drink respectively. This result maintains what is stated in studies done by (Izugbara et al., 2020).

Random Coefficient Model

It is possible to generalize the model so that the effect of level 1 covariates is different in each region. This can be done by adding random coefficients in front of some of the individual-level

covariates of the model. This model contains a random slope for wealth index and educational level, which means that it allows the effect of the coefficient of the explanatory variable to vary from region to region.

Table 5: Estimates of random coefficient model

Variables	Categories	β	S.E	Wald	sign	odds
Place of residence	Urban (Ref)	----	----	----	----	----
	Rural	0.534	0.128	4.13	0.000*	3.52
Age	15-24 (Ref)	----	----	----	----	----
	25-34	0.197	0.097	2.21	0.036*	1.130
	35-49	-0.384	0.151	-2.516	0.031*	0.843
Religion	Orthodox (Ref)	----	----	----	----	----
	Catholic	-0.451	0.321	-1.41	0.113	0.812
	Protestant	-0.356	0.232	-1.51	0.086	0.685
	Muslim	-0.615	0.498	-1.23	0.093	0.521
	Traditional	0.872	0.230	3.79	0.000*	1.97
	Other	0.575	0.327	1.751	0.071	1.23
Employment status	Not employed (Ref)	----	----	----	----	----
	Employed	-0.925	0.274	3.41	0.033*	0.54
Marital status	Single (Ref)	----	----	----	----	----
	Married	0.501	0.115	4.32	0.023*	2.071
	Separated	0.688	0.131	5.30	0.012*	3.252
Number of living children	0 (Ref)	----	----	----	----	----
	1-2	0.129	0.033	4.12	0.004*	2.52
	3-4	0.821	0.251	3.26	0.015*	1.64
	5+	0.932	0.315	2.95	0.037*	1.19
husband's alcohol consumption	Does not drink	----	----	----	----	----
	Drink/never get drunk	0.254	0.072	3.55	0.014*	1.256
	Gets drunk sometimes	0.681	0.126	5.452	0.001*	3.64
	Gets drunk very often	0.791	0.133	6.051	0.001*	5.19
Random effect						
	$\hat{\sigma}_{u_{\text{region}}}^2 = \text{Var}(u_{0j})$	0.293	0.195			
	$\hat{\sigma}_{u_{\text{zone}}}^2 = \text{Var}(u_{0j})$	1.652	0.983			
	$\hat{\sigma}_{1}^2 = \text{Var}(u_{1j})$	0.0282	0.0228			
	$\hat{\sigma}_{2}^2 = \text{Var}(u_{2j})$	0.0066	0.027			
	$\hat{\sigma}_{u1}^2 = \text{Cov}(u_{0j}, u_{1j})$	0.0075	0.0192			
	$\hat{\sigma}_{u2}^2 = \text{Cov}(u_{0j}, u_{2j})$	-0.0571	0.0415			
	$\hat{\sigma}_{12}^2 = \text{Cov}(u_{1j}, u_{2j})$	-0.0267	0.0397			
	ICC(region)	0.055				
	ICC(zone)	0.315				

Source: EDHS, 2016

The ICC increased and is estimated as $\rho^2 = 0.055$ by adding level 1 predictors. This indicated that 5.5% of the total variability in gender-based violence against women aged between 15 to 49 years is attributable to the random factor and region in the random coefficient multilevel binary logistic model. The random coefficient estimates for intercepts and the slopes vary significantly at a 5% significance level (Table 5). It implies that there is a considerable variation in the effects of educational level and wealth index; these variables differ significantly across the regions. The variance of intercept in the random slope model is 0.293, which is still large, relative to its standard error of 0.195. Thus, there remains some regional-level variance unaccounted for in the model. The variance corresponding to the slope of wealth is 0.0066, which is relatively small to its standard error; this suggests that the effect of the wealth index may be justified in constraining the effect to be fixed. Likewise, the variance corresponding to the slope of educational level is 0.0282, which is relatively large to its standard error (S.E = 0.0228); thus, this suggests that the effect of family education may be justified in constructing the effect to be random.

3.2. Testing for Spatial Dependency

Spatial dependency of Physical violence

Global moran I for regression residuals

```
## Weights: bristol.W
## Moran I statistic standard deviate=24.652, p-value<2.2e-14
## alternative hypothesis: greater
## sample estimates:
## observed Moran I      Expectation      Variance
0.324566895             -0.0024589654    0.000045623
```

The moran I test statistic is 24.652, and the p-value is under .05, which indicates that the researcher can reject the null hypothesis.

Lagrange Multiplier diagnostics for spatial dependence of Physical violence

```
## Weights: bristol.W
## LMlag: 12.365    df: 21,    p-value: 3.26e-9
```

Lagrange Multiplier diagnostics for spatial dependence of Physical violence

Weights: bristol.W

LMerr: 562.3, df: 21, p-value<2.36e-14

The LMlag statistic is 12.365, whilst the LMerr statistic is 562.3. Both have p-values under .05, indicating that both the spatial lag and error models would be preferred over the OLS model. Since both models are significant for physical violence, the researcher needs to do additional tests. That is, **Robust LM** tests, to identify tests which one could be at work.

Robust LM test For Physical violence

##Lagrange multiplier diagnostics for spatial dependency of Physical violence

weights: bristol.W

##RMLag = 0.06532545, df=21, p-value=0.564231

##Lagrange multiplier diagnostics for spatial dependency of Physical violence

weights: bristol.W

##RMerr = 503.145, df=21, p-value=3.2e-10

Therefore, we cannot reject the null hypothesis that there are spatial lags in the possible presence of spatial errors of Physical violence of women. So far the results suggest that a spatial errors model for Physical violence of women is the most appropriate type of model.

Spatial dependency of sexual violence

Global moran I for regression residuals

Weights: bristol.W

Moran I statistic standard deviate=44.254, p-value<1.3e-13

##alternative hypothesis: greater

sample estimates:

## observed Moran I	Expectation	Variance
0.546232152	-0.00365245892	0.004523156

The test statistic is 44.254, and the p-value is under .05, which indicates that the researcher can reject the null hypothesis.

Lagrange Multiplier diagnostics for spatial dependence of sexual violence

Weights: bristol.W

LMlag: 21.32654 df: 21, p-value: 2.45e-9

Lagrange Multiplier diagnostics for spatial dependence of sexual violence

Weights: bristol.W

LMerr: 456.4, df: 21, p-value<4.6e-24

The LMLag statistic is 21.326, whilst the LMerr statistic is 456.4. Both have p-values under .05, indicating that both the spatial lag and error models would be preferred over the OLS model. Since both models are significant for sexual violence, the researcher needs to do additional tests. That is, **Robust LM** tests, to identify tests which one could be at work.

Robust LM test for sexual violence

##Lagrange multiplier diagnostics for spatial dependency of sexual violence

weights: bristol.W

##RMLag = 0.098, df=21, p-value=0.56243

##Lagrange multiplier diagnostics for spatial dependency of sexual violence

weights: bristol.W

##RMerr = 6524, df=21, p-value=3.4e-12

Therefore, we cannot reject the null hypothesis that there are spatial lags in the possible presence of spatial errors of sexual violence of women. So far the results suggest that a spatial errors model for sexual violence of women is the most appropriate type of model.

Spatial dependency of emotional violence

Global moran I for regression residuals

Weights: bristol.W

Moran I statistic standard deviate=46.25, p-value<5.34e-9

##alternative hypothesis: greater

sample estimates:

## observed Moran I	Expectation	Variance
0.546232184	-0.0045623214	0.006652554

The test statistic is 46.25, and the p-value is under .05, which indicates that the researcher can reject the null hypothesis.

Lagrange Multiplier diagnostics for spatial dependence of emotional violence

Weights: bristol.W

LMlag: 56.213 df: 21, p-value: 4.46e-12

Lagrange Multiplier diagnostics for spatial dependence of emotional violence

Weights: bristol.W

LMerr: 401.9, df: 21, p-value<4.56e-9

The LMlag statistic is 56.213, whilst the LMerr statistic is 401.9. Both have p-values under .05, indicating that both the spatial lag and error models would be preferred over the OLS model. Since both models are significant for emotional violence, the researcher needs to do additional tests. That is, Robust LM tests, to identify tests which one could be at work.

Robust LM test for emotional violence

##Lagrange multiplier diagnostics for spatial dependency of emotional violence

weights: bristol.W

##RMLag = 0.0654231, df=21, p-value=0.12612

##Lagrange multiplier diagnostics for spatial dependency of emotional violence

weights: bristol.W

##RMerr = 129.31, df=21, p-value=3.8e-12

Therefore, we cannot reject the null hypothesis that there are spatial lags in the possible presence of spatial errors of emotional violence of women. So far the results suggest that a spatial errors model for emotional violence of women is the most appropriate type of model. Thus, for the three response variables, both simple tests of the lag and error are significant, indicating the presence of spatial dependence. There is a correlation between nine Ethiopian regional states, two administrative towns, and 68 administrative zones in terms of women's violence.

Spatial Autocorrelation

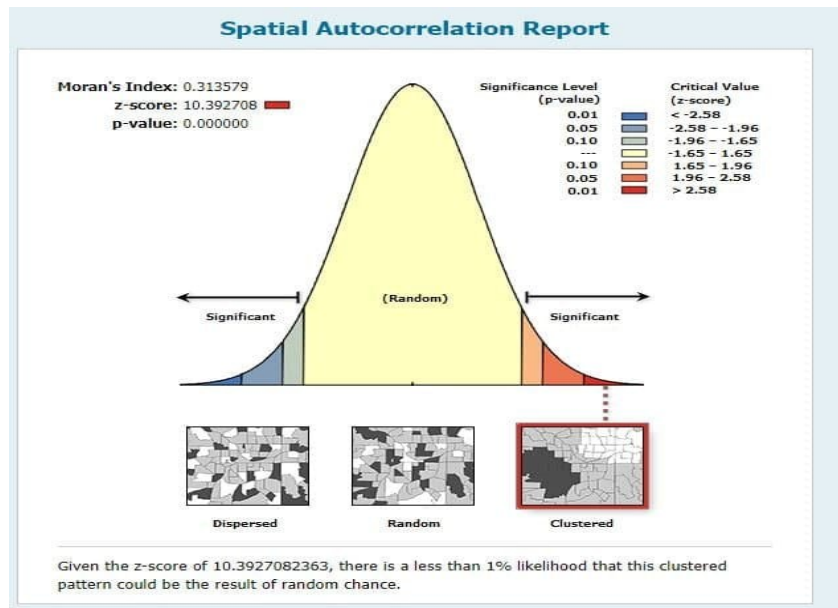


Figure 1: Spatial autocorrelation report of physical violence

Source: Ethiopian Demographic and Healthy Survey (EDHS, 2016)

The right side of each panel shows that a high rate of physical violence occurred over the study area. The auto-generated interpretations displayed underneath each panel show that the likelihood of clustered patterns occurring by chance is less than 1%. The dark red color indicates significant global clusters. Since the $p\text{-value} < 0.0001$ is statistically significant, the researcher rejects the null hypothesis. This indicated that there was a significant clustering of physical violence in Ethiopia. Therefore, the spatial distribution of high values and/or low values in the dataset was more spatially clustered than would be expected if underlying spatial processes were random.

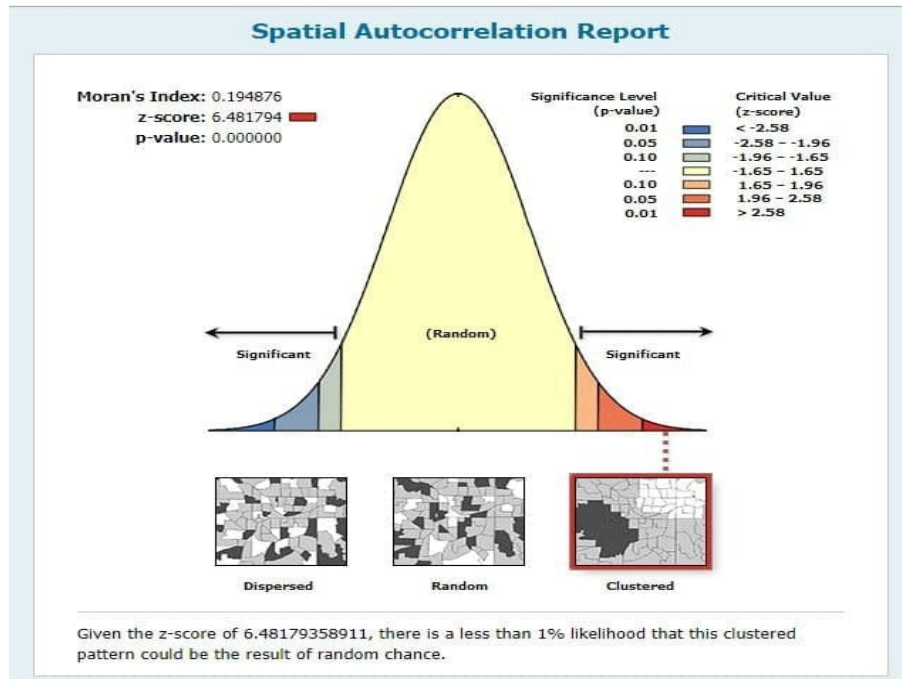


Figure 2: Spatial autocorrelation report of sexual violence

The right side of each panel shows that a high rate of sexual violence occurred over the study area. The auto-generated interpretations displayed underneath each panel show that the likelihood of clustered patterns occurring by chance is less than 1%. The dark red color indicates significant global clusters. Since the $p\text{-value} < 0.0001$ is statistically significant, the researcher rejects the null hypothesis. This indicated that there was a significant clustering of sexual violence in Ethiopia. Therefore, the spatial distribution of high values and/or low values in the dataset was more spatially clustered than would be expected if underlying spatial processes were random.

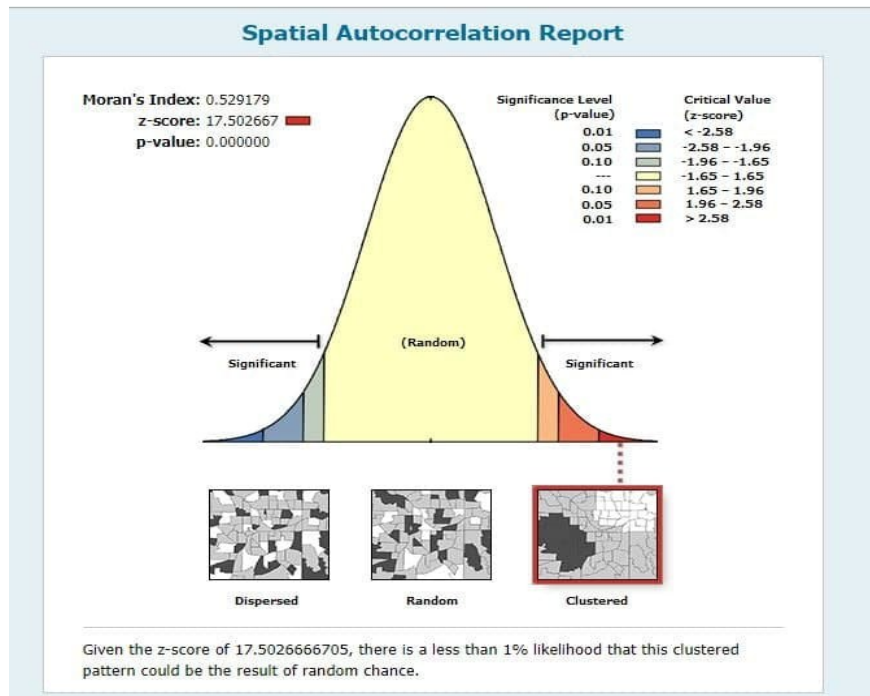


Figure 3: Spatial autocorrelation report of emotional violence

The right side of each panel shows that a high rate of emotional violence occurred over the study area. The auto-generated interpretations displayed underneath each panel show that the likelihood of clustered patterns occurring by chance is less than 1%. The dark red color indicates significant global clusters. Since the $p\text{-value} < 0.0001$ is statistically significant, the researcher rejects the null hypothesis. This indicated that there was a significant clustering of emotional violence in Ethiopia. Therefore, the spatial distribution of high values and/or low values in the dataset was more spatially clustered than would be expected if underlying spatial processes were random.

Hotspots analysis

a) Hotspots and Cold spots Analysis of physical violence in Ethiopia

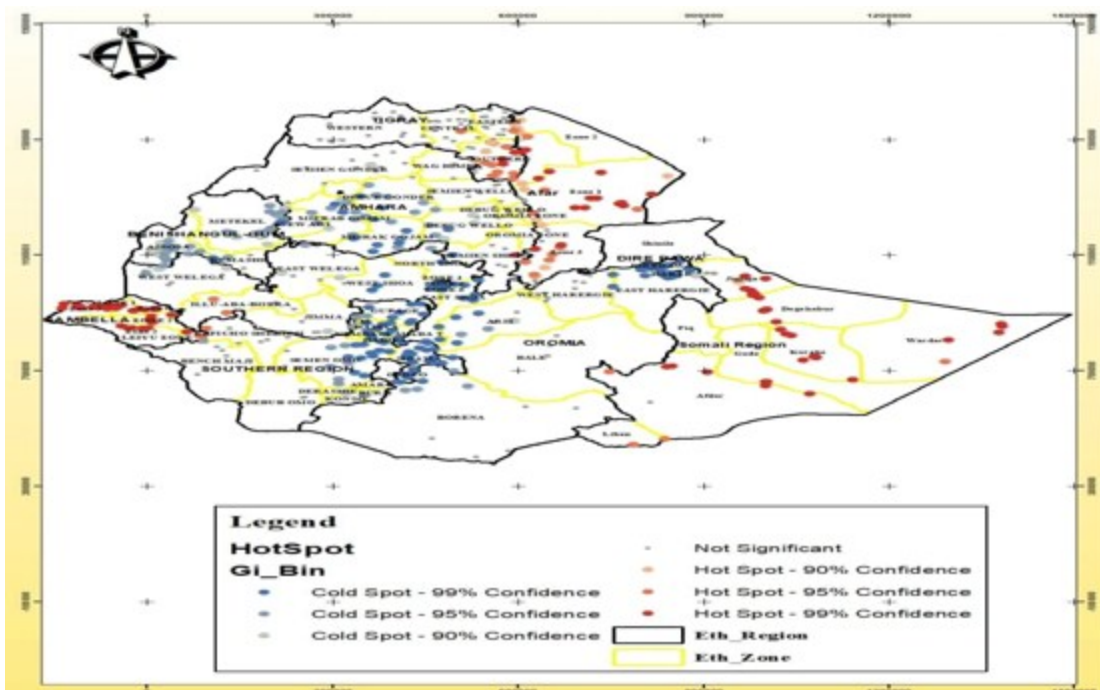
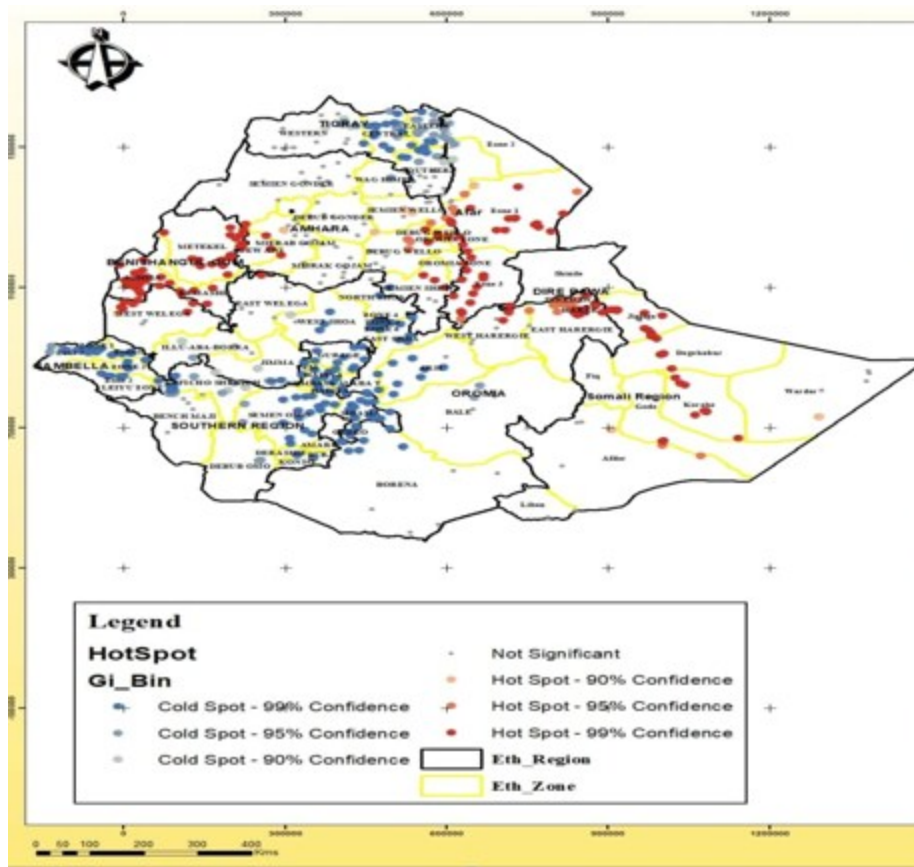


Figure 4: Hotspots and Cold spots Analysis of physical violence in Ethiopia

Source: Ethiopian Demographic and Healthy Survey (EDHS, 2016)

The red-brown color indicated the more intense clustering of significant hotspots (areas with higher rates) proportion of physical violence observed in the Gambella region especially (Anyuak Zone and Mejjeng Zone) and Afar region, especially around Gabi and Fanti Zone. In contrast, the blue color indicated areas with significantly lower rates of physical violence (cold spot areas) that were found in the Oromia region (Arsi and East Shewa Zone), some parts of SNNPR around Kambata Tambaro, Hadiya, and Gedeo Zone.

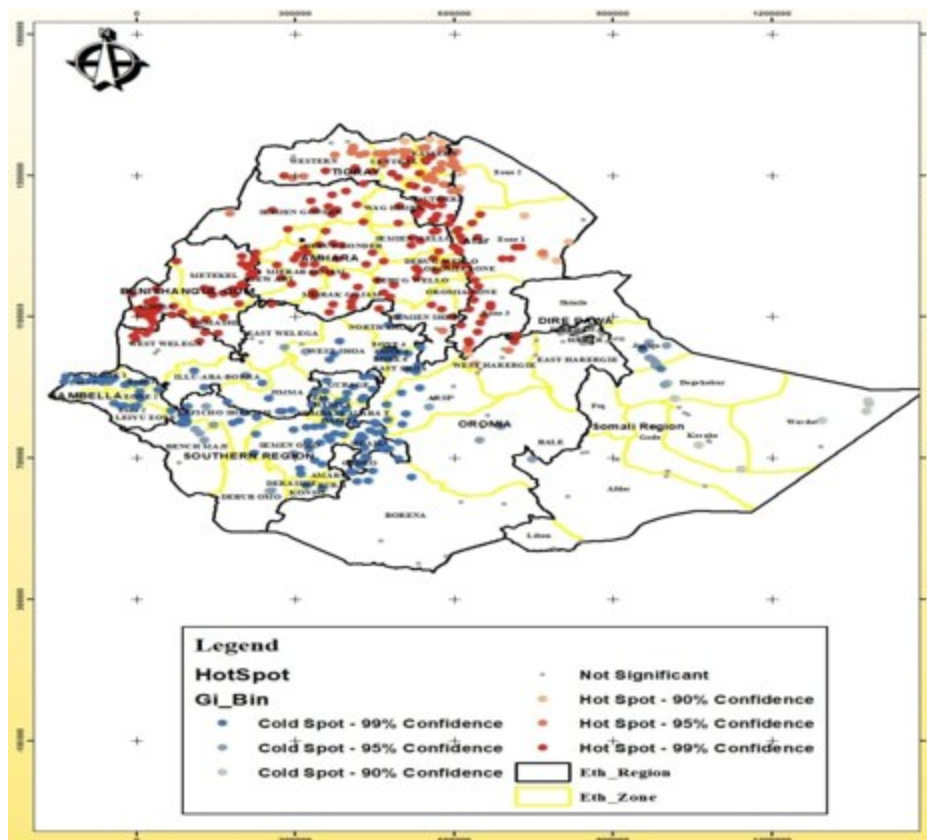
b) Hotspots and Cold spots Analysis of sexual violence in Ethiopia



Source: Ethiopian Demographic and Healthy Survey (EDHS, 2016)

The red-brown color indicated the more intense clustering of significant hotspots (areas with higher rates) proportion of sexual violence observed in the Benishangul-gumuz region mainly Metekel Zone and Kemashi Zone, and Afar region mainly Rasu Zone and Awsi Zone. In contrast, the blue color indicated areas with significantly lower rates of sexual violence (cold spot areas) that were found in the SNNPR around Semien Zone, Gedeo Zone, and Sidama Zone.

c) Hotspots and Cold spots Analysis of emotional violence in Ethiopia



Source: Ethiopian Demographic and Healthy Survey (EDHS, 2016)

The red-brown color indicated the more intense clustering of significant hotspots (areas with higher rates) proportion of sexual violence observed in the Amhara region around west Gojjam Zone, Awi Zone, and South Wollo Zone. In contrast, the blue color indicated areas with significantly lower rates of sexual violence (cold spot areas) that were found in the SNNPR around Gamo Zone, Gedeo Zone, and Wolayta Zone.

4. CONCLUSION

The results of this study showed that the norms and practices prevalent in male-dominated patriarchal societies, which are deliberately biased in favor of men and serve as a basis for justifying the continued use of domestic violence, particularly against women, have largely preserved power imbalances. Efforts to change these norms were often contradicted by male interests. In fact, in most developing countries, men dominate socio-political and economic activities, and even when gender equality laws are passed, they are practically neither defended nor implemented.

Additionally, religion and cultural beliefs promote domestic violence in Ethiopia. Therefore, in addition to the integration of various UN conventions on women's rights and gender equality, there is much-needed political will from senior governments, strengthened by binding administrative policies and guidelines. A legal framework should be put in place in Ethiopia to ensure that violations are dealt with appropriately. There is also a need for a socio-cultural reorientation so that our male-dominated societies understand that the mistreatment of women is barbaric, primitive and inhumane; that there are other dignified and kind ways to resolve domestic problems or conflicts than beating one's wife.

Finally, the lack of empowerment of women leads to domestic violence, women are considered inferior beings and are therefore victims of violence and there is no need to emancipate themselves economically and socially. Women who lived in such a cultural environment felt inferior and excluded when making decisions. Gender-based violence is also attributed to economic tensions, with economic instability leading to a decline in values and thus gender-based violence. The legal framework to combat gender-based violence against women has proven ineffective.

5. Recommendations

Based on the findings, the following recommendations were made:

- ✓ It is imperative for care providers to take into account the cultural background and distinct challenges encountered by victims of gender-based violence. They should also offer culturally competent services to children, youth, and families who have experienced gender-based violence.
- ✓ It is recommended that the state budget include a dedicated funding line for the implementation of measures aimed at combating gender-based violence, with a focus on thoroughly addressing the needs of survivors.

REFERENCES

- Abrahams N, Devries K, Watts C. (2014) Worldwide prevalence of non-partner sexual violence: a systematic review. *Lancet*; **383**: 1648–54
- Ayele G. and Kebedu T.(2020) Prevalence of intimate partner violence against women and associated factors in Ethiopia. *Womens Health*. 20:22
- Benebo F, Schumann B, Vaezghasemi M. (2018) Intimate partner violence against women in Nigeria: a multilevel study investigating the effect of women's status and community norms. *BMC Womens Health*.;18(1):136.
- C. O. Izugbara, M. O. Obiyan, T.T. Degfie and A. Bhatti, (2020), Correlates of intimate partner violence among urban women in sub-Saharan Africa. *PLoS ONE* 15(3):
- Central Statistical Agency (CSA, 2016); ICF. Ethiopia Demographic and Health Survey 2016; CSA: Addis Ababa, Ethiopia; ICF: Calverton, MA, USA,
- Cynthia O., Ph., Nuria S., Antor O. (2020) Perceived Determinants of Domestic Violence and the Strategies for Its Prevention in Orlu Local Government Area of Imo State, South East Nigeria.
- DeGue S, Valle L, Holt M, Matjasko J, Tharp A. (2014) A systematic review of primary prevention strategies for sexual violence perpetration. *Aggress Violent Behav*; **19**: 346–62.
- Devries K, Mak J, Bacchus L, et al. (2013) Intimate partner violence and incident depressive symptoms and suicide attempts: a systematic review of longitudinal studies. *PLoS Med*; **10**: e1001439
- Devries K, Mak J, García M. (2013) The global prevalence of intimate partner violence against women. *Global Health Science*; **340**: 1527–28.
- Eshetu A., Zeleke G., Yordanos S., and Yohannes Sh. (2021) Factors Associated with Attitude towards Wife-Beating among Married Women of the Reproductive Ages in Arba Minch Town, Southern Ethiopia: A Community-Based Cross-Sectional Study.
- Eunice J., Ines K., Emilia M., Mohsin S., and Kristien R. (2021) Domestic violence in Mozambique: from policy to practice.
- Fekadu E, Yigzaw G, Gelaye KA, Ayele TA, Minwuye T, Geneta T, Teshome DF. (2018) Prevalence of domestic violence and associated factors among pregnant women attending antenatal care service at University of Gondar Referral Hospital, Northwest Ethiopia. *BMC Womens Health*.;18(1):138
- Feldman J, Clifton D. (2014) Female genital mutilation/cutting: data and trends, update 2014. Washington, DC: Population Reference Bureau,
- Jean J., Joel K., Gertrude L., and Joseph M. (2019) Psychosocial Factors Associated with Domestic Violence Inflicted on Women by their Husbands at Kenge, Kwango Province, DRC.

Jepkoech, K. (2021). Determinants of Domestic Violence among Women Attending Gender Based Violence Centre (GBV) At Nakuru County Referral Hospital. *Journal of Medicine, Nursing & PublicHealth*, 4(2),19-31

Modi M, Palmer S, Armstrong A. (2014) The role of Violence Against Women Act in addressing intimate partner violence: a public health issue. *J Womens Health*; **23**: 253–59.

Muluken D., Virginia S., Lyn F. and Kingsley A. (2020) Gender Based Violence against Women in Sub-Saharan Africa: A Systematic Review and Meta-Analysis of Cross-Sectional Studies.

Ömer A. & Şeyda Ü. (2021)Determinants of Domestic Physical Violence against Women in Turkey

Rasha A. & Rania A. (2017) Magnitude and determinants of domestic violence against ever-married women in Sohag, Egypt

Stöckl H, Devries K, Rotstein A,. (2013)The global prevalence of intimate partner homicide: a systematic review. *Lancet*; **382**: 859–65.

World Bank (2014). Voice and agency: Empowering women and girls for shared prosperity. Washington, DC: World Bank Group,; 195

World Bank, (2019). Gender-based violence (Violence against women and girls). <https://www.worldbank.org/en/topic/socialsustainability/brief/violence-against-women-and-girls>.

World Health Organization (WHO, 2021).Violence against Women. <https://www.who.int/news-room/fact-sheets/detail/violence-against-women>.

Don W, Michael & Kristian S. (2008) *Spatial Regression Models*. Thousand Oaks, CA: SAGE. *Quantitative Applications in the Social Sciences*.

Coleman M., Coleman, Mi., Mabuza A., Kok, G., Coetzee M & Durrheim D. (2009). “Using the SaTScan method to detect local malaria clusters for guiding malaria control programs”. *Malaria journal*.

Filzmoser, P., Maronna, R., Werner, M. (2008). “Outlier identification in high dimensions.” *Comput Stat Data Analysis*.

Elhost, J.P. (2010). “Applied Spatial Econometrics”, *Raising the bar, Spatial Economic Analysis*.

Anselin & Bera. (1998). “Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics”. (1st Ed.). *CRC Press*.

Lesage, J. & Pace, R.K. (2009). "Introduction Spatial Econometrics", *London, England; Taylor & Francis*.

Catani, P.; Teräsvirta, T.; Yin, M. A. (2017). "Lagrange multiplier test for testing the adequacy of constant conditional correlation GARCH model". *Economy. Rev.*