*British Journal of Education, Society & Behavioural Science*
*16(4): 1-12, 2016, Article no.BJESBS.25899*
*ISSN: 2278-0998*

**SCIENCEDOMAIN** *international*
www.sciencedomain.org

# Statistical Modeling of Crime Severity in Sidama Zone and Its Neighbors, Ethiopia

**Aklilu Toma Shamenna[1*], Ashenafi Senbeta Bedane[1], Tilahun Ferede Asena[1,2] and Amanuel Disassa Abshoko[1]**

*[1]School of Mathematical and Statistical Sciences, Hawassa University, P.O.Box 05, Hawassa, Ethiopia.*
*[2]Department of Statistics, College of Natural and Computational Sciences, Arba Minch University, P.O.Box 25, Arba Minch, Ethiopia.*

***Authors' contributions***

*This work was carried out in collaboration between all four authors. Author ATS make substantial contributions to conception and design, and acquisition of data analysis and interpretation. Authors ATS, ASB, TFA and ADA did the statistical analysis and literature searches while analyses of study were by authors ATS and ASB. All authors participate in drafting the article and revising it critically for important intellectual content and all authors gave final approval of the version to be submitted and any revised version for publication.*

***Article Information***

DOI: 10.9734/BJESBS/2016/25899
Editor(s):
(1) Chan Shen, Department of Biostatistics, University of Texas, USA.
Reviewers:
(1) Alexandre Valença, Fluminense Federal University, Niteroi, Rio de Janeiro, Brazil.
(2) Nergis Canturk, Ankara University Institute of Forensic Sciences, Ankara, Turkey.
(3) S. Zimeras, University of the Aegean, Greece.
Complete Peer review History: http://sciencedomain.org/review-history/15137

***Original Research Article***

***Received 25th March 2016***
***Accepted 11th June 2016***
***Published 24th June 2016***

## ABSTRACT

**Background:** The study of crime and its determinants is closely related to factors that influence the mind and behavior of the individual in being participated in criminal activities.
**Objective:** The study has been undertaken with the aim to identify and analyze the factors that influence individuals in criminal activities Sidama Zone and its neighbors.
**Methods:** Multi stage Stratified random sampling with stratification formed by crime level was applied to this study. Proportional Odds Model has been employed to predict the Probability of dependent variable on the basis of independent variables and to model the effects of the covariates included in the study on crime severity of individuals.

**Corresponding author: E-mail: aketomy@gmail.com;*

**Results:** Age, sex, educational background, marital status, residence, employment status, partners education, attitude towards revenge, wealth status, awareness of criminal law, previous crime history and number of dependent families were found to be statistically significant risk factors.
**Conclusions:** It is recommended that concerned government bodies should have proper reporting system of crime. There should be better supply of education facilities to everyone to beware that crime is wrong practice. Government should provide employment opportunities; provide basic information regarding criminal law to every citizen so that crime will be under control.

*Keywords: Criminal activities; Ethiopia; proportional odds model; Sidama zone.*

## 1. INTRODUCTION

The study of crime and its determinants is closely related to poverty, social exclusion, income inequality, culture, family background, religion, unemployment, education, age, gender, race, urbanization and a host of other economic and socio-demographic factors that influence the mind and behavior of the individual in being participated in criminal activities [1]. As such, criminal activities are not restricted to economic systems. In addition, crimes are also not peculiar to levels of economic growth and development.

The circumstances surrounding the individual offender such as his/her personality, physical characteristics, intelligence, family background, environmental surrounding such as peer groups and neighbors have been subject of the study of crime [2]. Hence, understanding the attributes of criminals will be helpful to design and implement prudent crime prevention strategies.

### 1.1 Crime Definitions and Classification

As defined by [3] crime is an act or omission of an act, which is punishable by law. However, an act that is considered as a crime in one place and time may not be true in another place or time. Crime is classified in to two broad categories as **personal or violent crime** is any criminal offense which involves the use of or even the threat of force or violence and it has a broad legal category that encompasses a number of criminal offenses and **property crime** which includes those offences involving the loss of property during which there is no use of violence by the perpetrators.

### 1.2 Review of Related Literatures

It is suggested by [3] suggested that intent and opportunity are two major factors that lead to the occurrence of a crime. An individual cannot commit a crime unless and otherwise he/she gets an opportunity even though he/she is intended to commit a crime. Therefore, the best strategy for crime prevention is to provide a system that denies any opportunity for a criminal to commit a crime.

As seen by [1], it seems reasonable to expect that the level of property crime will be influenced in some way by the distribution of income (and wealth) while [4] argues that monetary and fiscal policies have impacts on crime.

Unemployment and crime relationship was tested by [5] using annual time series data in England and Wales over the period 1971-2000. It was evidenced in this study that in the long-run, the overall and individual property crimes are co-integrated with unemployment as well as law enforcement instruments. Particularly, unemployment rate has positive co-integration with overall crime, burglary and theft, indicating that, for such crimes, the motivation effect is stronger than the opportunity effect.

On the other hand [6] only finds support for the criminal motivation effect. Controlling for the variation in the unemployment crime relationship by age group and over time, he concluded that unemployment has a greater motivational effect on property crime among young adults. Controlling for police and race (black/white) confounders, he concluded that changes in inflation as well as changes in rates of cyclical and frictional unemployment are important predictors of property crime rates.

An eighteen-year-old is five times more likely to be arrested of a property crime than is a thirty-five-year-old; for violent crime the corresponding ratio is two to one and those aged fifteen to nineteen constituted roughly 7% of the overall population but accounted for over 20% of arrests for violent offenses and roughly one-third of all property-crime arrests. Males under age of eighteen are five times more likely to be arrested for violent crime than are females; for property crime the ratio is less than three to one. Violent

crime rises sharply in adolescence to a peak at age eighteen, before steadily declining thereafter. Property-crime offenses peak at age sixteen and drop off more quickly. The pattern for murders is similar to that of violent crime more generally [7].

The effect of unemployment on crime investigated by [8] for the period of 1988-1999 in Swedish countries by using fixed effect including time series and country level data. Variables are average income, education, social allowances population density considered. Linear and quadratic time trend techniques have been used. The result shows that unemployment has a positive and significant effect on crime.

The study in Argentina by [9] using panel data approach for the period of 1990-1999 in Argentina by incorporating GDP, unemployment rate, gini Coefficient, and other inequality variables revealed that the effect of unemployment rate and inequality rate is significantly associated with crime rate.

The study conducted by [10] investigated the relationship between unemployment, inflation and crime rate by using annual data for the period of 1970-2006 in Malaysia. Crime rate, inflation and unemployment rate variables are considered. Bartlett Corrected trace test technique has been used. The result shows that there is a positive significant effect of inflation and unemployment rate on crime but inflation is not positively associated with crime rate in short run.

Socio economic and demographic determinants of crime were investigated by [11] using panel data set for the period of 1993-1999 in Spain. Crime rate, education, GDP, growth and unemployment rate variables are considered. GMM estimation technique has been used. The result shows that the impact of socio-economic determinants on property crime while demographic variables are significantly associated with violent crime and also it indicated that there is positive significant effect of socio economic determinants on crime.

On the study conducted by [12], the relationship between unemployment and crime in sixteen states were investigated by using country level data. Variables are per capita income, age, population (black and white) and amount of federal funding for education considered. Microsoft excel was used to run separate regression analysis. The result shows that there is a significant effect between unemployment and crime.

The determinants of crime in the context of Pakistan were investigated by [12] using regression analysis on time series data for the period of 1980 to 2010 with the objective of finding out the determinants which are the main causes of increasing crime in Pakistan. The regression analysis shows that house hold consumption, GDP, population, literacy, wage rate and have strong positive significant effect on crime, while migrant in and out of Pakistan and electricity crisis have weak positive and significant on crime.

The relationship between and its various factors such as unemployment, poverty and inflation was analyzed by [13] by using time series data from 1975-2007 in Pakistan by using Granger causality test technique for determining the result. The result shows that crime has significant effect on unemployment, poverty, and inflation.

The effect of job loss on crime was investigated by [14] in Norway by using an exceptional individual panel data set with limitations of unmarried and employed Norwegian men under age of 40. Variables are plant closure, age, number of kids, committed crime, and income and net wealth. Specification test techniques have been used. The result shows that there is no significant effect of plant closure on violent crime, however a significant effect of plant closure on violent crime because of mental distress, frustration, and anger are the mechanism due to which job loss effect crime.

The linkage between inflation, unemployment and crime rate in Malaysia was investigated by [10] using annual data from 1970 to 2006. Inflation, unemployment and crime rate variables are considered. Bartlett corrected test technique have been used. The result shows that inflation and unemployment are positively related to the crime rate but inflation is not significant in short run.

In Ethiopia, crime statistics has shown that the rate of crime is increasing steadily. There are a number of factors which may be responsible for the growing rate of crime such as unemployment, economic backwardness, over population, illiteracy and inadequate equipment of the police force [3].

The activities that person do may be un common to the norms, culture, values and traditions for the others, and that may be reason for occurrences of crimes. Crime is a complex social phenomenon and its cost is increasing due to a number of societal changes and the like, and hence law enforcement organizations like that of police need to learn the factors that contribute to the high crime trends [15].

There is a need for crime prevention strategies and policies and therefore the study of factors that cause the increment of crime is important for crime prevention and control as discussed above in different literatures. Accordingly, this study focuses on identifying factors that explain the rate and occurrence of crime in Ethiopia. The results of study will help as guiding tool in design and implementation of crime prevention strategies in Ethiopia and will be baseline for further studies of crime.

## 2. DATA AND METHODOLOGY

### 2.1 Description of the Study Area and Population

The study was conducted in Sidamma zone and its two neighboring zones namely Gedeo zone and West Arsi zone in Ethiopia.

### 2.2 Sidama Zone

Sidama is one of the most populous zones of Southern Nations Nationalities and Peoples Regional State of Ethiopia. According to the recent census [16], the total population of the zone was 2,954,136. With an area of 6,538 square kilometers, Sidama has a population density of 452/km2 with an average household size of 4.99 persons. From the total population, 5.51% are urban inhabitants and 0.18% are pastoralists.

### 2.3 Gedeo Zone

Gedeo zone is a Zone in the Southern Nations, Nationalities and Peoples Region (SNNPR) of Ethiopia. This Zone is named for the Gedeo people, whose homelands lie in this zone. Gedeo extends south as a narrow strip of land along the eastern escarpment of the Ethiopian Highlands into the Oromia Region, which borders the Zone on the east, south and west; Gedeo shares its northern boundary with Sidama zone.

The 2007 census conducted by the central statistical agency of Ethiopia reports this Zone has a total population of 847,434, of which 424,742 are men and 422,692 women. From the total population, 12.72% are urban inhabitants [16]. A total of 179,677 households were counted in this Zone, which results in an average of 4.72 persons to a household.

### 2.4 West Arsi Zone

West Arsi zone is one of newly established zone since May 2008 in Oromia regional state of Ethiopia. Based on socio-economic profile, the zone is inhabited by a total population of 1,844,542 of which 85.27% population is considered as rural, and the remaining 14.73% is classified as urban. The female population accounts for 50.5% of the urban and 49.5% of the rural population [17]. According to census conducted by central statistical agency, the household size is 7.2 for rural and 5.6 for urban. This is higher than the country level average household size of 3.9 for urban and 4.9 for rural, and for Oromiya Regional State 3.8 in urban and 5 in rural [16].

### 2.5 Sampling Procedure

Sampling methods are scientific procedures of selecting those sampling units that would provide the required estimator with associated margins of uncertainty arising from examining only a part not the whole of the population. Multi stage Stratified sampling with stratification formed by crime level was applied to this study. The main purpose of stratification is to reduce sampling error. Moreover, stratified sampling is a technique which uses any relevant information that might be available, in order to increase efficiency. It involves the division or stratification of a population by partitioning the sampling frame in to non-overlapping and relatively homogeneous groups.

### 2.6 Sample Size Determination

In the planning of a sample survey or researches, a stage at which a decision must be made about the size of the sample is always required. However, too large a sample implies wastage of resources, and too small a sample diminishes the utility of the results [18]. Therefore, the decision should be made with a minimum cost but the estimate will explain the population characteristics with a high precision.

However, several formulas developed for sample size calculations that conform to different research situations. The sample size for this study was determined based on stratified sampling at a 95% confidence level using general formula for sample size determination adopted [11] as:

$$n = \frac{\sum_{h=1}^{3} \frac{W_h^2 S_h^2}{w_h}}{V + \frac{1}{N}\sum_{h=1}^{3} W_h S_h^2} \quad (1)$$

Let Z be the upper α/2 point of standard normal distribution, where α=0.05 so that $Z_{0.05}$=1.96.

The total population N=13,835 from 3 crime severity level and 3 sampled area are selected for primary data which contain $N_1$ (Sidama Zone)=4,869, $N_2$ (Gedeo Zone)=3946 and $N_3$ (West Arsi Zone)=5020. This strata total is again allocated for three crime levels in the study. After all, the proportional allocation using $W_h \frac{N_h}{N}$ and $S_h{}^2$ are more convenient for computing the sample size from the first minimum estimated sample size$n_0$.

$$n_o = \frac{\sum_{h=1}^{3} W_h S_h^2}{V} = \left(\frac{Z_{\alpha/2}}{Nd^2}\right)^2 \sum_{h=1}^{3} W_h S_h^2 = 1259$$

Based on the above results, we found the following sample sizes for three level of crime (serious, medium and low) and location (Sidama Zone, Gedeo Zone and West Arsi Zone) as proportional allocation.

$$n_h = \left(\frac{N_h}{N}\right) * n, \quad h = 1,2,3$$

| Crime categories /zones | Series | Medium | Low | Total |
|---|---|---|---|---|
| Sidama zone | 739 | 352 | 244 | 1308 |
| Gedeo zone | 559 | 320 | 378 | 1257 |
| West Arsi zone | 447 | 341 | 480 | 1268 |
| Total | 1745 | 986 | 1102 | 3833 |

## 2.7 Description of Data and Methods of Data Collection

Crimes are recorded by police on daily basis. This study is based on both primary data and secondary data. The questionnaire consisting poverty, social exclusion, income inequality, cultural and family background, religion, unemployment, education, age, gender, race, urbanization and a host of other economic, socio-demographic and environmental factors that influence the mind and behavior of the individual in making decision in participating in crime activity was designed. The primary data was collected by using a designed questionnaire and arrested persons were asked to complete the modified Amharic version questionnaire; while the secondary data of respondents which was obtained from record of police commotion offices using identification number.

## 2.8 Variables Considered under the Study

The response variable in this study is crime severity categorized as **serious crime** includes crimes such as murder, corruption, crimes against public institutions and the like, **medium crimes** includes rape, vandalism, fraud, and crimes against individual personality and **low crimes** such as smuggling trade, drug use and transmission and so on. There are forty three types of crime in the cities and all those types are classified in to three categories (serious, medium and low). The classification of crime is performed by regional law enforcement agencies. The categorization has taken in to account minimum number of years the offender is penalized according to different types of crimes.

The most important and common factors (covariates) that influence the incidences of crime which are obtained based on literatures are age, sex, educational background, marital status, residence, employment status, wealth status, partners education, attitude towards revenge, awareness of rule of law, previous crime history and number of dependent families.

## 2.9 Method of Data Analysis

In order to meet the objective set up on this study Ordinal logistic regression model and tests related are employed as a general methodology.

## 2.10 Ordinal Logistic Regression Model

Logistic regression model can be classified as multinomial, ordinal and binary. In this investigation, ordinal logistic regression model

was used. The ordinal logistic regression procedure empowers one to select the predictive model for ordered dependent variable. It describes the relationship of an ordered response variable and a set of explanatory variables. The explanatory variables may be continuous or discrete (or any type). The most popular model in ordinal logistic model is the proportional odds model which is described below.

## 2.11 Proportional Odds (PO) Model

In the proportional odds model, the cumulative logits can be represented as parallel linear function of independent variables. That is, for each cumulative logit, the parameters of the models are the same except for the intercept. Consequently, according to the proportional odds assumption, odds ratio is the same for all categories of the response variable. Let **Y** takes categorical response variable with **c** ordered categories and assume *pro( Y=1) is p1, pro( Y=2) is p2,……, pro(Y= i) is pi for i=1,2,……,c.*

Cumulative probability reflect the ordering, with *pro( Y≤1) ≤ pro( Y≤2) ≤ ,……. ≤ pro(Y≤i) =1* and let the cumulative probability of the first **c -1** of Y is pro(Y≤ *j*) = πi for i=1,2,… c-1. Then the odds of the first **c-1** cumulative probability are:

$$Odds\left(\Pr ob\left(Y \le i\right)\right) = \frac{\Pr ob(Y \le i)}{1 - \Pr ob\left(Y \le 1\right)} = \frac{\pi_i}{1 - \pi_i}, \quad i = 1,\ 2, ...., c-1 \qquad (2)$$

In the Proportional Odds (PO) model, the log odds of the first c -1 cumulative probabilities are given as:

$$\log it\left[pro\left(Y \le i\right)\right] = \left[\frac{pro\left(Y \le i\right)}{1 - pro\left(Y \le i\right)}\right] = \log\left[\frac{\pi_i}{1 - \pi_i}\right] \qquad (3)$$

And the relationship between the cumulative logits of **Y** is:

$$\log\left[\frac{\pi_i}{1 - \pi_i}\right] = \log\left[\frac{\pi_i}{\pi_{i+1} + .... + \pi_c}\right]; i = 1, 2, ..., c-1 \qquad (4)$$

Consider the collection of p explanatory variables for the $j^{th}$ subject denoted by the vector $X_j^T = (x_{1j}, x_{2j}, ....., x_{pj})$, j=1,2,….., n. In logistic regression, the relationship between the predictor and response variables is the logit transformation of $\pi_j(X_j)$.

$$\pi_i\left(X_j\right) = \frac{\exp\left(\alpha_i - \beta_1 x_{11} + ... + \beta_p x_{pj}\right)}{1 + \exp\left(\alpha_i - \beta_1 x_{11} + ... + \beta_p x_{pj}\right)}$$
$$for\ i = 1,2,.....c-1,\ j = 1,2,....., n \qquad (5)$$

Where **β** is a vector of **p** regression coefficients and $\alpha_i$ is the $i^{th}$ intercept coefficient.

Then the logit or log-odds of having pro(Y≤ i) =πi (the first i cumulative probability) is modeled as a linear function of the explanatory variables as:

$$\log\left[\frac{pro\left(Y \le i\right)}{1 - pro\left(Y \le i\right)}\right] = \log\left[\frac{\pi_i}{1 - \pi_i}\right] = \alpha_i - \beta_1 X_{1j} - ... - \beta_p X_{pj} \qquad (6)$$

Therefore,

$$\log it\left[pro\left(Y \le i\right)\right] = \alpha_i - \sum_{i=1}^{p} \beta_i X_{ij}; \quad 0 \le \pi_i \le 1;$$
$$i = 1, 2, ..., c-1,\ \ j = 1, 2, ..., n \qquad (7)$$

The model assumes a linear relationship for each logit and parallel regression lines. Equation (7) is called proportional odds model and it estimates simultaneously multiple equations of cumulative probability.

## 2.12 Parameter Estimation of Ordinal Logistic Regression Model

The ordinal logistic regression model is fitted to the observed responses using the maximum likelihood approach that is, by maximizing the likelihood, or more usually, by maximizing the logarithm of the likelihood (loglikelihood). The likelihood function is given by the equation:

$$L\left(\beta^*\right) = \prod_{i=1}^{n} \prod_{j=1}^{c-1} \left[\frac{\exp\left(\alpha_i - X_j^{'}\beta\right)}{1 + \exp\left(\alpha_i - X_j^{'}\beta\right)}\right]^{y_{ij}} \left[1 - \frac{\exp\left(\alpha_i - X_j^{'}\beta\right)}{1 + \exp\left(\alpha_i - X_j^{'}\beta\right)}\right]^{1 - y_{ij}} \qquad (8)$$

Here **β*** a vector containing both the slope and intercept coefficients.

It follows that the log-likelihood function is

$$\ln\left[L\left(\beta^*\right)\right]=\sum_{i=1}^{n}\sum_{j=1}^{c-1}\left\{y_{ij}\ln\left[\frac{\exp\left(\alpha_i - X'_j\beta\right)}{1+\exp\left(\alpha_i - X'_j\beta\right)}\right]+\left(1-y_{ij}\right)\left[\frac{\exp\left(\alpha_i - X'_j\beta\right)}{1+\exp\left(\alpha_i - X'_j\beta\right)}\right]\right\} \quad (9)$$

The estimates of the unknown parameters are obtained by differentiating the log-likelihood function with respect to each parameters and setting each result to zero [19].

### 2.13 Model Selection and Adequacy Checking Criteria

It is much better to compare models based on their results, reasonableness and fit as measured for example by the Akaike Information Criterion (AIC) and Bayessian Information Criteria (BIC). The AIC and BIC criteria can be used to compare the stability of competing models. AIC ans BIC are defined as:

$$AIC = -2L + 2P$$
$$BIC = -2\log(L) + p\log(n)$$

Where L is log likelihood of the fitted model, P is numbers of parameters in the model in the fitted model and n is the number of observations. From a set of competing models, the best model is the one with lowest value of AIC and BIC [20].

### 2.14 Deviance

The deviance measures can also be used to check the adequacy of the fitted model. The deviance is defined as

$$D = -2\left[\ln(L_s) - \ln(L_f)\right]$$

Where Ls is the log-likelihood of the simpler model, Lf is the log-likelihood of the full model **k** is the degree of freedom which is the difference in the number of parameters between the full model and the saturated model.

### 2.15 Evaluating the Logistic Regression Model

After fitting the multiple logistic regression model using the selected predictor variables, the next step is to assess the fit of the model to the data.

For the selected model we should look at an overall test of the null hypothesis that the location coefficients for all of the variables in the model are zero before proceeding to examine the individual coefficients. The change in likelihood function has a chi-square distribution even when there are cells with small observed and predicted counts. This value provides a measure of how well the model fits the data. The log likelihood statistic is analogous to the error sum of squares in multiple linear regressions.

### 2.16 Goodness of Fit Measures

The structural form of the model describes the patterns of association and interaction. The sizes of the model parameters determine the strength and importance of the effects. Model's predicted values smooth the data and provide improved estimates of the mean of Y at possible explanatory variable values.

The deviance is also used to construct a goodness of fit test for the model. The goodness of fit statistics for ordinal logistic regression has a form:

$$D = 2\sum\sum O_{ij}\log\left[\frac{O_{ij}}{E_{ij}}\right]$$

Likewise, the Pearson chi-square statistic also compares the model fit to the actual data, defined

$$X^2 = \sum\sum\frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$$

Where $O_{ij}$ is the observed value and $E_{ij}$ is the expected value for the $i^{th}$ observation. Both goodness-of-fit statistics should be used only for models that have reasonably large expected values in each cell.

### 2.17 Testing of the Parallel Lines

The basic assumptions under the implementation of proportional odds model is the assumption of parallel lines. That is, it is assumed that the slope coefficients in the model are the same across the

response categories. A non-significant test of parallelism is taken as evidence that the logit surfaces are parallel and that the odds ratios can be interpreted as constant across all possible cut points of the outcome variable. The chi-square test statistic is used for testing the hypothesis of parallelism and is defined as the difference between of -**2loglikelihood** of the constrained model (the proportional odds model with **c-1** intercept and **p** slopes and the general model which fits separate **c** category multinomial models to each of the **p** predictors).

### 2.18 Statistical Test of Individual Parameters

The statistical significance of individual regression coefficients is tested using the Wald statistic which is the squared ratio of the estimated logistic coefficients to their standard error [19].

$$Z^2 = \left[\frac{\hat{\beta}_j}{se(\hat{\beta}_j)}\right]^2 \quad (10)$$

## 3. RESULTS OF THE STUDY

The main objective of the study is to identify and analyze the factors that influence individuals in criminal activities in Sidama Zone and its neighbors in Ethiopia. The analysis has been performed for 3833 individuals from three zones and from three crime levels. Statistical computations were carried with STATA software.

### 3.1 Ordinal Logistic Regression Analysis

For the purpose of selecting candidate predictors for the multiple ordinal logistic regression model, fifteen uni-variable ordinal logistic regression model were considered. From the results, it was evidenced that the variables employment type (p-value=0.203), marriage type (p-value=0.612) and satisfaction with working environment (p-values=0.199) of respondents are insignificant factors among the variables considered under the study. Therefore, we consider two models: Model I containing all the fifteen covariates while Model II excludes the three insignificant covariates from the analysis. The following table displays the results of goodness of fit tests for Model I and Model II.

From the Table 1, we see that the deviance based chi square value is equal to 1532.505 (p-value=0.000) and 1543.932 (p-values=0.000) for Model I and Model II respectively. This is an indication that both Model I and Model II fit the data well. Moreover, the computed values for the Pearson and deviance test statistics were found to be insignificant for the two models. This is a further indication that the two models fit the data well.

Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) were among the most commonly used statistics in the selection of best model among two or more candidate models. Here our focus is on how to use AIC and BIC statistics to select one of the two models as our final model. The 2log likelihood for Model I and Model II are 14,881.98 and 14,768.091 respectively. Having this, the computed value of the deviance test statistic becomes:

$$D = -2[\ln(L_I) - \ln(L_{II})] = 14{,}88198 - 14{,}768091 = 113889$$

From the chi-square distribution table we obtain that the percentage value of chi-square distribution with two degrees of freedom at 1% level of significance is equal to 9.21. Since the deviance test statistic is much greater than this tabulated value of chi-square distribution, we conclude that the deviance test statistic is insignificant, and this in turn indicates that Model II is a better fit to our data.

To check the proportional odds assumption, we use parallel line test as discussed in the methodology part. The following table displays the results of parallel lines assumption tests of Model II.

**Table 1. Results of deviance based tests for model adequacy**

| | Model | -2Log likelihood | Chi-square | df | Sig |
|---|---|---|---|---|---|
| Model I | Intercept only | 16,525.912 | | | |
| | Final | 14,881.98 | 1643.932 | 34 | 0.000 |
| Model II | Intercept only | 16,300.596 | | | |
| | Final | 14,768.091 | 1532.505 | 30 | 0.000 |

**Table 2. Parallel lines test assumption for model II**

| | -2Log likelihood | Chi-square | df | Sig |
|---|---|---|---|---|
| Null hypothesis | 14,768.091 | | | |
| General | 14,714.689 | 53.402 | 30 | 0.103 |

**Table 3. Parameter estimates for significant covariates in the final model**

| Variable | Categories | Est. | SE | Wald | df | Sig. | 95% CI | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | LB | UB |
| Crime type | Series | -1.704 | 0.128 | 177.3 | 1 | 0.000* | -1.955 | -1.453 |
| | Medium | -0.693 | 0.108 | 41.18 | 1 | 0.000* | -0.905 | -0.481 |
| Age | 18-30 | -0.556 | 0.098 | 32.201 | 1 | 0.01 | -0.575 | -0.364 |
| | 31-50 | -0.466 | 0.071 | 43.078 | 1 | 0.001 | -0.605 | -0.327 |
| | Above 50 (ref) | - | - | - | - | - | - | - |
| Sex | Male | 0.153 | 0.049 | 9.75 | 1 | 0.007 | 0.057 | 0.249 |
| | Female (ref) | - | - | - | - | - | - | - |
| Educational background | Illiterate | -2.72 | 0.656 | 17.167 | 1 | 0.000* | -4.006 | -1.433 |
| | Primary | -3.549 | 0.667 | 28.323 | 1 | 0.000* | -4.858 | -2.242 |
| | 2ndary & above (ref) | - | - | - | - | - | - | - |
| Marital status | Single | -1.225 | 0.0922 | 176.53 | 1 | 0.000* | -1.436 | -1.044 |
| | Married | -1.023 | 0.057 | 322.11 | 1 | 0.000* | -1.133 | -0.912 |
| | Widowed | -0.855 | 0.065 | 173.02 | 1 | 0.000* | -0.981 | -0.728 |
| | Divorced (ref) | - | - | - | - | - | - | - |
| Residence | Rural | -1.142 | 0.059 | 373.34 | 1 | 0.000* | -1.255 | -1.026 |
| | Urban (ref) | - | - | - | - | - | - | - |
| Employment status | Unemployed | -0.325 | 0.022 | 218.23 | 1 | 0.000* | -0.368 | -0.282 |
| | Employed (ref) | - | - | - | - | - | - | - |
| Partners education | Illiterate | -0.397 | 0.118 | 11.389 | 1 | 0.001* | -0.627 | -0.166 |
| | Primary | -0.469 | 0.193 | 5.913 | 1 | 0.015 | -0.848 | -0.091 |
| | 2ndary and above (ref) | - | - | - | - | - | - | - |
| Attitude towards revenge | Negative | -1.134 | 0.25 | 20.57 | 1 | 0.0041 | -1.624 | -0.644 |
| | Positive (ref) | - | - | - | - | - | - | - |
| Wealth status | Poor | -0.244 | 0.33 | 54.676 | 1 | 0.000* | -0.891 | 0.402 |
| | Medium | 0.069 | 0.022 | 9.81 | 1 | 0.002 | 0.026 | 0.112 |
| | Rich (ref) | - | - | - | - | - | - | - |
| Awareness of rule of law | Yes | 0.506 | 0.101 | 25.127 | 1 | 0.000* | 0.308 | 0.704 |
| | No (ref) | - | - | - | - | - | - | - |
| Previous crime history | Never committed | 0.602 | 0.111 | 29.61 | 1 | 0.002 | 0.384 | 0.82 |
| | Committed & not arrested | 0.072 | 0.023 | 10.04 | 1 | 0.000* | 0.027 | 0.117 |
| | Committed & arrested (ref) | - | - | - | - | - | - | - |
| Number of dependent families | No dependent family | -2.476 | 0.23 | 116.315 | 1 | 0.000* | -2.927 | -2.025 |
| | 1-3 | -0.14 | 0.023 | 38.15 | 1 | 0.000* | -1.185 | -0.095 |
| | Above 3 (ref) | - | - | - | - | - | - | - |

*Ref- reference category, *significant at 1% level*

From the results of tests of parallel lines in the (Table 2), we can see that the chi-square test statistic is not significant.

Therefore, we have no enough evidence to reject the null hypothesis stating that the slope coefficients in the model are the same across the response categories. Thus, the proportional odds assumption appears to hold for the final model (Model II).

The Table 3 displays the results of estimates of significant covariates of the final multiple ordinal logistic regression model.

## 4. DISCUSSION

The study investigates socio-economic, demographic and demographic factors that lead individuals to participate in criminal activities.

One of the greatest problems facing modern society is crime. Pervasive crime robs us of our sense of safety and security, causing psychological harm even to those who are never actually victimized. It deprives victims of their property or even their lives. It threatens the stability of our families. It costs society a tremendous amount of money. Coping with crime is one of the most important functions of local, state, and national governments. To solve the problem of crime, we must attack its causes. Different studies indicated that poverty is the major cause of crime [11,12,13]. From our study results, it was indicated that wealth status is one of the significant covariates for crime. The result indicated that individuals with poor category of wealth status are at higher risk to commit crime than individuals who are under rich category of wealth status. The adjusted odds ratio [exp(0.244)=1.276] indicates that poor individuals are 1.276 times more likely to be involved in series crimes as compared to rich individuals. The odds ratio could be as low as 0.147 and as high as 2.42. This result is in agreement with the findings of [11,12].

The relationship between aging and criminal activity has been noted since the beginning of criminology. The proportion of population involved in crime tends to peak in adolescence and then decline with age [21]. It is now a truism that age is one of the strongest factors associated with criminal behavior. In fact some studies have claimed that the age-crime relationship is universal across groups, societies and times and that this universality signals that the age-crime relationship is strongly biologically determined [22]. From our study, the results indicated that age is significant covariate of crime. The adjusted odds ratio [exp(0.556)=1.51] indicates that individuals under the age category of 18-30 years are 1.51 times more likely to commit series crimes as compared to those individuals under the age category of above 50 years. This result is also in agreement with the findings of [21].

The study conducted in Punjab by [23] showed shows that education, family size, income level and land dispute have strong significant effect on crime. One of the covariates used in our study is education status which is categorized as illiterate, primary, secondary and above. The results of ordinal logistic regression in this case indicated that illiterate individuals are at higher risk to commit crime than educated individuals. This result agrees with the result obtained in Punjab by [23].

The study showed that the odds of committing high crime are higher for individuals who are unmarried as compared to those who are married. This result agrees with previous study by [14] in which the effect of job loss on crime was investigated by using an exceptional individual panel data set with limitations of unmarried and employed Norwegian men under age of 40.

We also obtained that among the covariates included in our crime study, residence is a significant covariate. The adjusted odds ratio (aOR) [exp(1.142)=3.13] indicates that individuals from rural areas are 3.13 times more likely to commit serious crimes compared to individuals from urban areas. The odds ratio could be as low as 2.789 and as high as 3.51 with 95 % confidence. The result is consistent with the result found in Brazil by [24].

The current study showed that employment status is also significant covariate. The adjusted odds ratio [exp(0.325)=1.384] indicates that unemployed individuals are 1.384 times more likely to commit series crimes as compared to those individuals who are employed. The odds ratio could be as low as 1.33 and as high as 1.445 with 95% confidence. The result is consistent with previous studies on crime, on which the incidence of crime is higher among unemployed individuals than those who are employed by [4,10,11,13,14].

One of the social factors included in this study is partner's education and it significantly explains

the crime rate in the region. Moreover, previous crime history is significantly related with the dependent variable (current crime). That is, individuals' history of committing crime before they were arrested for committing current crime was one of significant covariate. Compared to those individuals who committed crime and arrested previously, those individuals who never committed crime in their life history before committing the current one are 1.83 times more likely to commit series crimes.

## 5. CONCLUSION AND IMPLICATIONS OF THE STUDY

The current study revealed that age, sex, educational background, marital status, residence, employment status, partner's education, attitude towards revenge, wealth status, awareness of criminal law, previous crime history and number of dependent families are the most important determinants of crime severity in Sidama zone and its neighbors in particular and in Ethiopia in general. Specifically, high crimes are more likely among individuals from rural areas as compared to those from urban areas.

Individuals with negative attitude on revenge are at higher risk to commit high crimes than those with positive attitude on revenge. Individuals with no dependent family are at higher risk of committing high crimes than those with dependent families. It has been also observed that high crime is more likely to be committed among individuals whose partners are not educated.

From the results of our study, we see that individuals from rural areas are more likely to commit crimes as compared to those individuals from urban areas. This may be because of awareness problem among societies living in rural areas regarding crime and its consequences. Therefore, we recommend for the concerned government agencies to work on awareness creation regarding crime and its consequences among societies living in rural areas via mass medias and public discussions.

The concerned government bodies should have proper reporting system of crime. There should be better supply of education facilities to everyone to beware that crime is wrong practice. Government should provide employment opportunities; provide basic information regarding criminal law to every citizen so that crime will be under control.

Different literatures indicated that the distribution of crime has spatial pattern [25,26,27]. Even though our study considered three different geographical areas, the space-time nature of the dependent variable is not taken in to consideration. Therefore, further studies are recommended by taking space-time nature of the dependent variable in to consideration.

## ACKNOWLEDGEMENTS

Author's sincere thanks go to Hawassa University vice president for research and technology transfer directorate for providing the financial support.

## COMPETING INTERESTS

Hereby all the authors have no competing interests.

## REFERENCES

1. Madden P, Chiu WH. Burglary and Income Inequality. Journal of Public Economics. 1998;69:123-141.
2. Andargachew T. The crime problem and its correction. Volume I Addis Ababa; 1988.
3. Thakur C. Crime control; 2003. Available:http://ncthakur.itgo.com/chand3c.htm
4. Trogdon H. The unemployment - crime relationship; 2006.
5. Lu Han. Economic analyses of crime in England and Wales; 2009.
6. Britt CL. Reconsidering the unemployment and crime relationship. Variation by age group and historical period. Journal of Quantitative Criminology. 1997;13(4):405-428.
7. Grönqvist H. Youth unemployment and crime: New lessons exploring longitudinal register data; 2011.
8. Edmark K. The effects of unemployment on property crime: Evidence from a period of unusually large swings in the business cycle; 2003.
9. Maria Cerro A, Meloni O. Determinants of crime rate in Argentina during 90's. 2004; *27*.
10. Foon Tang C. The linkages among inflation, unemployment and crime rates in Malaysia. Int. Journal of Economics and Management. 2009; 3:1.
11. Buonanno P, Montolio D. Identifying the socioeconomic determinants of crime across Spanish provinces; 2005.

12. Dr. Aurangzeb. Determinants of crime in Pakistan. Universal Journal of Management and Social Sciences. 2012; 2:9.
13. Gillani SY, Rehman HU, Gill AR. Unemployment, poverty, inflation and crime nexus: Cointegration and causality analysis of Pakistan; 2009.
14. Regea M, Skardhamar T, Telle K, Votruba M. The effect of job loss on crime: Evidence from plant closure events; 2009.
15. Wilson JQ. Crime and public policy. In Wilson and Petersilia, eds. Crime: Public policies for crime control. Institute for Contemporary Studies. 2004;537-538.
16. CSA. Summary and statistical report of the 2007 population and housing census results. Addis Ababa; 2007.
17. West-Arsi zone ARDO. Socio-economic Probfile of West-Arsi zone, Shashamane, Ethiopia; 2008.
18. Cochran WG. Sampling techniques. John Wiley & Sons, New York, 3rd Edition; 1977.
19. Hosmer Diw, Lemarshow S. Applied logistic regression. John and Sons, Inc; 1989.
20. Agresti A. Categorical data analysis. John Wiley & Sons, New York, 2nd Edition; 2002.
21. Quetelet A. Research on the property for crime at different ages (S. Sylvester, Trans). Cincinnati, OH: Anderson (Original work published 1931); 1984.
22. Kanzawa S, Still MC. Why men commit crimes (and why they dessist). Sociological Theory. 2000;18:434-447.
23. Mehmood K, Cheema MA. Determinants and maximum likelihood functions of juvenile crime in Punjab, Pakistan. International Journal of Agriculture & Biology; 2008.
24. Hartung G, Pessoa S. Demographic factors as determinants of crime rates; 2007.
25. Anselin L, Cohen J, Cook D, Gorr W, Tita G. Spatial analyses of crime, criminal justice, V4. Measurement and Analysis of Crime and Justice, National Institute of Justice, Washington D.C; 2000.
26. Andresen MA, Brantingham PJ. Hot Spots of crime in Vancouver and their relationship with population characteristics. Department of Justice Canada. Ottawa; 2007.
27. Crutchfield R, Geerken M, Gove W. Crime Rates and social integration: The impact of metropolitan mobility. Criminology. 1982; 20:467-478.



*Peer-review history:*
*The peer review history for this paper can be accessed here:*
*http://sciencedomain.org/review-history/15137*