INVESTIGATING FACTORS AFFECTING RATE OF VIOLENT CRIME IN KENYA USING MULTIPLE LINEAR REGRESSION

BY

THOMAS OMYONGA

REG.NO. I56/68658/2013

M.S.C. SOCIAL STATISTICS

SHOOL OF MATHEMATICS

COLLEGE OF BIOLOGICAL AND PHYSICAL SCIENCES

UNIVERSITY OF NAIROBI

A research project submitted in partial fulfillment of the requirement for the degree of Masters of Science in Social Statistics

June 2015

DECLARATION

CANDIDATE

This project is my original work and has not been presented for a degree in any other University.

Signature ……..Date ……

Thomas Omyonga

Reg. no. I56/68658/2013

SUPERVISOR

This project has been submitted for examination with my approval as supervisor

Signature……..Date……..

Prof. Moses M. Manene

SHOOL OF MATHEMATICS

COLLEGE OF BIOLOGICAL AND PHYSICAL SCIENCES

UNIVERSITY OF NAIROBI

P.O.BOX 30197-00100

NAIROBI

DEDICATION

I dedicate this work to none other than my God in heaven. All the Glory be to Him.

ACKNOWLEDGEMENT

I take this opportunity to thank Almighty God for taking me through this journey towards getting this prestigious award of master's degree. To Him alone I owe it all. I acknowledge the Kenya National Bureau of Statistics for providing data for this study. I extend my appreciation to my supervisor Professor Manene for his great encouragement and guidance. My sincere gratitude to the people that I love most; my wife Josephine and my daughters; Ashley and Stacey. I thank you for your great support and encouragement. I also acknowledge my dear brother Ronald and his wife Sheila for their unwavering support and encouragement. All my friends and prayer partners; Etene,Steve,Zack,Chris,Keya,Were,Chacha,Mulamula,Pastor David among others ,I am grateful to you all. Finally to all my lecturers and the UON community, not forgetting Dr Were J. who is my mentor, I thank you so much and may the good lord bless you and your families mightily. For you have been angels to me. Thank you all.

ACRONYMS AND ABBREVIATIONS.

HDR	Human Development Report
SRIC	Security Research Information Centre
ENRL	Enrollment
ICPC	International Centre for the Prevention of Crime
GDP	Gross Domestic Product
GINI	A standard economic measure of income inequality based on Lorenzo curve. It is named after its inventor Corrado Gini
w.r.t.	With respect to
SSC	Secondary School Enrolment
FIR	Fire Arms and Ammunitions
POL	Police Ratio
GP	Gross Domestic Product
DAB	Drug Abuse
CNV	Conviction Rate
EPR	Employment Rate

ABSTRACT

The purpose of this study was to investigate the key factors affecting rate of violent crimes in Kenya. Data used in this study was obtained from the Kenya Bureau of Statistics. Exponential Population growth model was used to estimate annual Population. I used multiple linear regression model and R Computer software package to analyze the data. In this research study I identified the following violent crimes in Kenya: Murder, Manslaughter, Rape, Assault, Break-ins and Robbery. Others include: defilement, domestic violence, tribal clashes, carjacking, cattle rustling and terrorism. I then classified the crimes into two main categories in line with the Kenyan law. These are: Crime Against Person that includes Murder, Assault, Grievous harm, etc and Crime related to Property which includes Theft, Robbery, Arson, etc. My research has established that the following are key factors affecting violent crime rate: Police ratio, Number of fire arms and ammunitions recovered or surrendered, and conviction rate.

TABLE OF CONTENTS Page

LIST OF TABLES AND FIGURES Page

CHAPTER ONE: INTRODUCTION

1.1 Background

Van der Westhuizen (1977) defines crime in the following two ways : in juridical sense, it refers to a violation of the law for which the state (courts) may impose a punishment. In criminological terms, it includes all forms of criminal acts as well as deviant behaviour that are not necessarily regarded juridical in nature. In this research study i have relied on the juridical definition of crime.

Gimode (2001) says that crime is as old as human society. This is born out of the fact that, at all times in history, societies have had taboos, customs and laws designed to contain crime. He defines violent crime as the physical criminal violence or physical assault that includes the following crimes: homicide, armed robbery, car-jacking, attempted murder, manslaughter, rape, etc. lt involves the physical or psychological injury to other people.

Albert (1998) says that there can be no pursuit of meaningful sustainable development in a situation marked by uncontrolled and uncontrollable crime, because it breeds and accompanies insecurity. When the rate of violent crime is high, then the level of human insecurity is also high and vice versa.

Human Security was first popularized by the United Nations Development Program in the early 1990s. It emerged in the post-Cold War era as a way to link various humanitarian, economic, and social issues in order to alleviate human suffering and assure security. Supporters of this broad definition of human security as outlined in the 1994 Human Development Report(HDR) argue that hunger, disease, pollution, affronts to human dignity, threats to livelihoods, and other harms in addition to violence should all be considered human security issues.

However, supporters of the narrow definition of human security argue for a focus on violent threats to individuals and communities. They argue that human insecurity is the combination of threats associated with war, genocide, and the displacement of populations. At a minimum, it refers to the threat of violence and the fear of violence, Human security research group (2010).

In Kenya, there is a growing concern across the nation about the heightened incidences of criminal and violent acts. This means that the rate of violent crime is ranked high among

problems facing Kenyans. Reported cases of victims of these acts are Kenyan citizens of all cadre and from across the political divide.

In this research study I have modeled the data that was available to me using Multiple linear regression. I have determined the key determinants of violent crime rate and established how each affects the rate of violent crime. I believe that the results of this study will be of great help to the people of this country and other stake holders.

1.2 Problem statement

The problem of human insecurity is well known to Kenya. This problem has been studied in varied dimensions particularly in line with the 1994 Human Development Report(HDR).According to the HDR 1994,organized crime and criminal violence, human rights and good governance, armed conflict and intervention, genocide and mass crimes, health and development, resources and development, etc, are all determinants of human insecurity .

However, very few researchers have studied the rate of violent crime as the main indicator of human insecurity. In this research project I have investigated factors affecting rate of violent crime and established how each factor affects violent crime rate in Kenya.

1.3 Research Objectives

The main objective was to investigate key factors affecting rate of violent crimes in Kenya using multiple linear regression.

Specific objectives were:

- Investigate all identified factors based on available data
- Establish a Mathematical relationship between the factors and crime rate
- Identify statistically significant factors

1.4 Significance of the study

Albert (1998) says that there can be no pursuit of meaningful sustainable development in a situation marked by uncontrolled and uncontrollable crime, because it breeds and accompanies insecurity. Thus a high rate of violent crime if not controlled will breed a serious human insecurity problem which can divert the country from its optimal development path in addition to loss of lives of innocent citizens.

The findings of my research study will inform the government and other stake holders on key factors contributing to the rate of violent crime and the effect of each factor. This will facilitate the finding of a lasting solution to the problem of high violent crime rate. It will also inform other scientists interested in further research on the same problem.

The next chapter focuses on the literature review. It is a selective review of theoretical and empirical contributions to the scientific literature dealing with criminal behavior and its impact on human insecurity. Chapter III of my research paper focuses on the research design used for the project, how the factors are modeled including the statistical model to be used. Chapter IV focuses on data analysis that is it presents the results of regression analysis and how they are related to the research hypotheses. Finally chapter V indicates the major conclusions of my study and policy implications derived from my findings.

CHAPTER TWO: LITERATURE REVIEW

2.1 Violent Crime, Insecurity and development

Gimode (2001) conducted a research on the anatomy of violent crime and insecurity in Kenya. Taking a case study of Nairobi, the researcher established that crime disturbs the security of a society and hence provokes a situation of insecurity. Gimode (2001) describes violent crime as the Physical criminal violence or physical assault that includes homicide, armed robbery, car-jacking, attempted murder, manslaughter, rape, house breaking, etc. Chesnais (1992) agree with Gimode(2001) on the fact that violent crime has a direct impact on the state of human security in a society.Chesnais (1992) argues that violent crime is a big contributory factor that disturbs security in society.

Albert (1998) agrees with both Gimode (2001) and Chesnais (1992) on the issues of crime, insecurity and development. Albert (1998), there can be no pursuit of meaningful sustainable development in a situation marked by uncontrolled and uncontrollable crime, because it breeds and accompanies insecurity. A crime monitoring report by the security research and information centre (SRIC), Kenya agrees with all the three researchers, Gimode (2001), Chesnain (1992) and Albert (1998).The report is a result of a research project undertaken by SRIC on behalf of the government of Kenya. The main objective of the project was to strengthen the capacity to measure and monitor the incidences and impact of crime and armed violence in the country. Based on this objective, SRIC undertook to monitor crime incidences through capturing a wide array of crime incidences from mass media platforms, namely newspapers, published reports and social media. This exercise commenced in the month of July 2012 through June 2013. The report states that Security is one of the key basic human needs. Without it, the society is paralyzed and it cannot move. The report adds that for any country to develop economically the crime levels ought to be very low.

2.2 Weapons Used to commit Crime

In criminal cases where firearms were used, it was found that Assault rifles and Pistols were the preferred choices according to the SRIC report of June2013. Other weapons used included bombs, crude weapons and explosives. Indeed, a total of 293 Assault rifles, 106 pistols, 18 bombs, 134 crude weapons and 13 explosives were used to commit crimes during the reporting period of July 2012 to June 2013 (SRIC report, June2013).

Kirui and Mwaruvie (2012) conducted research on the effect of hosting refugees and physical insecurity in North-Eastern Kenya. The researchers established that sophisticated weaponry had found their way into Garisa district promoting banditry, cattle rustling and general violence in the district. This was attributed to the high influx of refugees from neighbouring countries into North- Eastern Kenya. Apart from refugees, several pastoral communities in Kenya, especially those in the border districts have been involved in the proliferation of small arms and lethal weapons such as the ones already mentioned above (Kirui and Mwaruvie, 2012).

According to the report by the International Centre for Prevention of Crime (2010), firearms are used for 40% of all homicides committed throughout the world. According to data processed by International Crime Victim Surveys, firearms are frequently used for thefts, robberies or attempted robberies and other violations, the report adds. Firearms are also used frequently for sexual crimes in Africa Van Kersteren (2003).

2.3 Measurement of Crime

Measures of crime include simple counts of offences, victimizations or apprehensions, as well as population based crime rates. Counts are normally made over a year long reporting period. This is according to an article on crime statistics Howit (2009). As regards the counting rules, the writer says that there are relatively few known standards that exist and out of these, there is none that permit international comparability beyond a very limited range of offences. However, many jurisdictions accept the following:

- There must be a prima facie case that an offence has been committed before it is recorded. That is either police find evidence of an offence or receive a believable allegation of an offense being committed. Some jurisdictions count offending only when certain processes

happen, such as an arrest is made, ticket issued, charges laid in Court or only upon securing a conviction.

- Multiple reports of the same offence usually count as one offence. Some jurisdictions count each report separately; others count each victim of an offence separately.
- Where several offences are committed at the same time, in one act of offending; only the most serious offence is counted. Some jurisdictions record and count each and every offence separately, others count cases, or offenders, that can be prosecuted.
- Where multiple offenders are involved in the same act of offending only one act is counted when counting offences but each offender is counted when apprehended.
- Offending is counted at the time it comes to the attention of a law enforcement officer. Some jurisdictions record and count offending at the time it occurs.
- In some countries, "only causing pain" is counted as assault but not in others.

Generally, offending that is a breach of the law but for which no punishment exists is often not counted, the writer adds. For example: Suicide, which is technically illegal in most countries, may not be counted as a crime, although attempted suicide and assisting suicide are. The writer states that because crime is a social issue, comparisons of crime between places or years are normally performed on some sort of population basis.

The writer lists the following methods of measuring crime among others;

Household surveys, hospital or insurance records, and compilations by police and similar law enforcement agencies. According to the writer, Public surveys are sometimes conducted to estimate the amount of crime not reported to police.

2.4 Factors Contributing to Crime

Oliver'02, Alison (2002), The Economics of Crime: An Analysis of Crime rates in America" based his theoretical framework on Becker's (1968) Rational Choice Model. In this model, Becker (1968) argues that an individual committing crime faces costs from law-enforcement activities. He adds that the severity of the punishment including fines and jail time is one part of the total cost, and the other part is the probability of getting caught. Becker (1968) concludes that Crime reduction can occur through reducing the benefits of crime, raising the probability of being caught, or increasing the costs of punishment conditional upon being caught. In addition to

Becker (1968) model, Isaac Ehrlich (1973) considers the effect of both Income and unemployment on crime rates. He argues that payoffs to crime, especially property crime, depend primarily on the opportunities provided by potential victims of crime as measured by the median income of the families in a given community. He also considers the effect of unemployment on crime rates. He views the unemployment rate in a community as a complementary indicator of income opportunities available in the legal labor market. Oliver'02, Alison (2002). The level and growth of economic activity in society create attractive opportunities for employment in the legal sector, but since they also improve the wealth of other members of society, the size of the potential loot from crime also rises.

Another factor Ehrlich (1973) discusses, which relates to the effect of economic conditions on the opportunity cost of crime, is the level of education of the population. Education helps to determine the expected rewards from both legal and criminal activities. Also, as a nation's average educational attainment increases, income will increase and the payoff of property crimes increases, thus increasing criminal's benefits. It is possible however, that school enrollment (ENRL) alone will reduce the time available for participating in the crime industry, and therefore have an inverse relationship on crime rates. (Ehrlich 1973).

Oliver'02, Alison (2002) says that together with the relationship between economic conditions and crime, one of the main issues in the pioneering studies of Becker (1968) and Ehrlich (1973) is the assessment of the effects of police presence, convictions, and the severity of punishments on the level of criminal activity. This is because people who are considering whether to commit crimes or not are assumed to evaluate both the risk of being caught and the associated punishment. Oliver'02, Alison (2002), Criminologists view the change in the population age distribution as the greatest influence on crime trends. As a general rule, the crime rate follows the proportion of young males in the population. Thus, the age(percentage of population under age 25) variable is expected to have a direct relationship with the crime rate.

Several other researchers agree with Becker (1968) and Ehrlich (1973) on the issue of classification of factors that are significant in determining the crime rate patterns over time .In addition to these, we have Biological factors such as heredity and brain activity, and also

Hormonal effects for example testosterone homone.Others include social factors such as Peer influence, Drugs and Alcohol, Child Abuse or Neglect(bad Child care),Home security such as security guards, and Guard dogs and also Easy access to Fire Arms (Lombroso, Cesare 1968,Curran,Daniel J. and Claire M.Renzetti 2001,Bowly,John A.,1988Arrigo,Bruce A.,1999,etc.

2.5 Crime Reporting, Recording and Classification

There are wide variations in recording and reporting practices, for certain offences and social groups, and in the classification of offences according to International Centre for the Prevention of Crime, ICPC (2010) report .This report agrees with Ernst Freund (1915) in his journal on Classification and definition of crimes. There are Legislation differences across all nations in the world and as a result, the ingredients of offences vary between jurisdictions Ernst Freund (1915). Ernst Freund (1915) says that crime statistics recording practices vary, not only between countries and jurisdictions but sometimes within jurisdictions and even between two individual law enforcement officers encountering the same situation. As regards classification of criminal offences, Ernst Freund (1915) states that in order to measure crime in a consistent manner, different sorts of crime need to be classified and separated into groups of similar or comparable offences.

The International Crime victims Survey has been done in over 70 countries to date and has become the 'de facto' standard for defining common crimes. Ernst Freund (1915)

In Kenya, violent criminal offences are classified as follows:

-Offences Against person include Assault, grievous harm, murder, etc.

-Offences Related to property include Theft, robbery with violence, arson, etc. (Kenya Economic Survey, 2004, 2009 and 2014)

2.6 Government Initiatives in response to Crime and Human Insecurity

Insecurity has become a perennial problem in Kenya. New cases of murder and crimes related to terrorism are always emerging in new locations in the country. Methods employed by the criminals are also changing with every new case. This has kept the security arms of the government on their toes as they try to contain the situation. In view of this, the government has

pursued a myriad of policy regimes legislative and legal frameworks and programmatic interventions aimed at reducing crime and human insecurity. Key among them is the Police reforms that are being implemented gradually and systematically keeping in line with the constitution and the Kenya Police strategic plan 2003-2007. This includes the annual Police recruitment of new officers.

Government facilitation to the Police service has increased. This is through provision of new equipment and access to new technology by Police officers through appropriate training. The process of Vetting Police officers is in progress and is aimed at improving efficiency in service delivery. The government is also committed to establishing more Police posts in different parts of the country as has been witnessed in Tunyai market, Tharaka South District among others. More Police officers and more undercover intelligence officers are deployed to areas that are badly affected by high crime rate as the crimes evolve, SRIC Report (June 2013). However, despite all these efforts, the high rates of violent crimes pause a serious negative impact on human security in the country.

2.7 Summary of the Literature Review

The aforementioned studies have highlighted several factors that are significant in determining the crime rate patterns over time. These factors have been classified into Economic, Deterrence, Demographic, Biological and Social variables. Economic variables include Education measured as the percentage of the population enrolled in secondary school, Gross Domestic Product (GDP) per capita, GINI measured as the Gini Index and Unemployment rate measured as the percentage of the population unemployed. Deterrence variables include the Police Ratio measured as the number of Police per capita, conviction rate measured as the number of prisoners convicted per capita. Demographic variables include Age measured as the percentage of the population under age 25 years. Biological variables include heredity and brain activity and also homones. Social variables include Peer influence, Drugs and Alcohol abuse, bad Child Care, Home security and Easy Access to Fire Arms. In the next chapter, a multiple linear regression model is used to model some of these factors.

CHAPTER THREE: METHODOLOGY

This chapter will focus on the research design used for the project, how the factors will be modeled and the statistical model to be used.

3.1 Data

In this research study, I used secondary data that is continuous for all the variables. I obtained the data from Kenya Bureau of Statistics. A total of 13,430 households were randomly selected from 1,343 clusters and surveyed. All surveyed households were captured using the Geographical Positioning System (GPS). The surveys were conducted over periods of 12 months each time. Both diary and recall methods were used in collecting household consumption and purchase information. The survey captured data that was used to update poverty, welfare statistics and employment statistics, derive the consumer price index, and revise the national accounts information. It also provided data on socio-economic aspects of the Kenyan population including education, health, energy, housing, water and sanitation. In addition to the survey, crimes data was obtained from the Kenya Police records. To obtain data on Population estimates, I applied the exponential Population growth model on the results of the 1999 and 2009 Population censuses.

3.2 Human Population growth

Definition:

Annual population growth rate for year t is the exponential rate of growth of midyear population from year t-1 to t, expressed as a percentage. Population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship except for refugees not permanently settled in the country of asylum, who are generally considered part of the population of the country of origin.

3.3 The Exponential Growth Model

This model is commonly called the **natural growth model.**

3.3.1Model Assumptions:

For this model we assume that the population grows at a rate that is proportional to itself.

The size of a Population growing exponentially at any time t, can be modeled as

$N_t = N_0 e^{tr}$ Where:

N_t = the number of individuals alive at time t,

N_0 = the number of individuals alive at time t = 0.

r = the exponential rate of increase of the Population.

In my data analysis, I used this model and MS Excel to:

(i) Determine the value of r

(ii) estimate population between the year 1999 and 2009 and

(iii) project the population between the year 2009 and 2013 using same value of r.

3.4 Variables in the model:

This study used the identified variables from the literature review supported by the available data. These are;

- Percentage of the population enrolled for Secondary education
- The Gross Domestic Product (GDP) per 100,000.
- Employment Rate measured as the percentage of the population employed.
- Police Ratio measured as the number of Police per 100,000.
- Conviction rate measured as the percentage of population convicted of violent crimes.
- The percentage of the population reported to be in possession of dangerous drugs.
- The number of Fire Arms and Ammunitions seized or surrendered per 100,000 of the population.

Due to unavailability of data, Employment rate was used instead of Unemployment rate.

Therefore, the variables being tested are:-

Dependent variable (Y):- Crime rate per year measured as the number of crimes reported to the Police per 100,000 of Population.

Independent variables:-

- Percentage of the population enrolled for Secondary Education
- The Gross Domestic Product
- Employment rate
- Police Ratio
- Conviction Rate
- The percentage of the Population reported to be in possession of dangerous drugs.
- The number of Fire Arms and Ammunitions seized or surrendered.

3.5 Calculation of Rates and Percentages

1. Crime Rate per 100,000 inhabitants: This represents the number of crime counts per 100,000 inhabitants.

I.e. crime rate per 100,000 = $\frac{No.of\ crimes}{Total\ population} \times 100{,}000\ \%$

2. Percentage secondary school enrolment = $\frac{\boldsymbol{No.of\ students\ enroled\ in\ secondary\ school}}{\boldsymbol{Total\ Population}} \times \mathbf{100\ \%}$

3. GDP per 100,000. = $\boldsymbol{GDP\ per\ capita \times 100{,}000}$

4. Employment Rate = $\frac{\boldsymbol{Total\ Number\ of\ employees}}{\boldsymbol{Total\ Population}} \times \mathbf{100\ \%}$

5. Number of Police personnel per 100,000 inhabitants (Police ratio):

$= \frac{\boldsymbol{Total\ number\ of\ Police}}{\boldsymbol{Total\ Population}} \times \mathbf{100{,}000}$

6. Percentage of Population convicted (conviction rate)

$= \frac{\boldsymbol{Number\ convicted\ of\ violent\ crime}}{\boldsymbol{Total\ Population}} \times \mathbf{100\ \%}$

7. Percentage of Population in possession of dangerous drugs

$= \frac{\boldsymbol{Total\ number\ in\ possession\ of\ drugs \times 100}}{\boldsymbol{Total\ Population}}\ \mathbf{\%}$

8. Number of fire arms and ammunitions per 100,000

$$= \frac{\boldsymbol{No.of\ fire\ arms}\&ammunitions}{\boldsymbol{Total\ Population}} \times \mathbf{100,000}$$

3.6 Linear Regression

Linear regression attempts to model the relationship between two variables by fitting a linear equation to observed data. One variable is considered to be an explanatory variable, and the other is considered to be a dependent variable.

A linear regression line has an equation of the form $y = a + bx$ where x is the explanatory variable and y is the dependent variable. The slope of the line is b and a is the intercept (i.e. the value of y when x = 0). For a linear relationship, we use a model of the form:

$y = \beta_0 + \beta_1 x_1 + \varepsilon,$ where ε is the error term in the model representing random fluctuations, measurement errors, or the effect of factors outside control.

3.6.1 Multiple Linear Regression Model

The response variable is often influenced by more than one predictor variable. A linear model relating the response y to several predictors has the form:

$Y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_1 \ + \cdots + \ \beta_k x_k + \varepsilon_i$

The parametersβ_0,β_1,…..β_k, are called regression coefficients.

3.6.2 Assumptions of Multiple Linear Regression:

Given the model $Y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \ \beta_k x_k + \varepsilon_i$,

(i) E($Y_i/x_1, x_2, x_{3,} ,,,,x_n$) = $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \ \beta_k x_k$ (conditional expectation of y given the predictor variables), E(ε_i) = 0 (i.e. the model is appropriate, relation is linear and all error terms have expected value zero)

(ii) Gaussian noise : ε_i ~N[$0, \sigma^2$] i.e. the error terms are approximately normally distributed about a mean of zero and variance σ^2

(iii) Constant variance; all the error terms have a common variance.

(iv) Independence of errors; all the error terms are independent.

3.6.3 Least-Squares Regression

The method of least – squares calculates the best-fitting line for the observed data by minimizing the sum of the squares of the vertical deviations from each data point to the line. To fit the line, we find the optimal values of the $\beta_i's$ i.e. we minimize the equation;

$\sum_{i=1}^{n} \{y_i - (\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k)\}$ ^2 = Q

We determine partial derivatives of Q w.r.t β_i; i = 0, 1, 2, ----,k

i.e. $-2\sum_{i=1}^{n}\{ y_i - (\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k)\} = \frac{\partial Q}{\partial \beta_0}$

$-2\sum_{i=1}^{n}\{x_1[y_i - (\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k)]\} = \frac{\partial Q}{\partial \beta_1}$

$-2\sum_{i=1}^{n}\{x_k[y_i - (\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k)]\} = \frac{\partial Q}{\partial \beta_k}$

We get the k+1 normal equations:

	$y_i \quad \beta_0 \quad \beta_1 x_1 \quad \beta_2 x_2 \quad \beta_3 x_3 \quad - - - - \quad \beta_k x_k$
1	$\sum y_i = n\beta_0 + \beta_1 \sum x_1 + \beta_1 \sum x_2 + \beta_3 \sum x_3 + - - - + \beta_k \sum x_k$
x_1	$\sum x_1 y_i = \beta_0 \sum x_1 + \beta_1 \sum x_1^2 + \beta_2 \sum x_1 x_2 + \beta_3 \sum x_1 x_3 + - - + \beta_k \sum x_1 x_k$
-	- - - - --- --- -- -- -- -- --- -- -- -
x_k	$\sum x_k y_i = \beta_0 \sum x_k + \beta_1 \sum x_k x_1 + \beta_2 \sum x_k x_2 + \beta_3 \sum x_k x_3 + - - + \beta_k \sum x_k^2$

In matrix form the model is; $\boldsymbol{Y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ and the estimators are found by solving the matrix equation

$\boldsymbol{\beta} = (X^1X)^{-1}X^1\mathbf{Y}$ where $\boldsymbol{\beta} = \begin{pmatrix} \beta 0 \\ \beta 1 \\ \beta 2 \\ \beta k \end{pmatrix}$

Hypothesis:

$H_0 : \beta_i = 0$ i.e. variable X_i has no predictive value Vs H_1: $\beta_i \neq 0$ i.e.X_i has added predictive value.

Test Statistics:

$t = \frac{\boldsymbol{\beta_i}}{se\sqrt{\frac{n\,c_{ii}}{n-k-1}}} = \frac{\boldsymbol{\beta_i}}{se\beta_i}$ Where c_{ii} = the ii^{th} element of the matrix

$$(X^1X)^{-1} = \begin{pmatrix} n & \sum X_1 & \sum X_k \\ \sum X_1 & \sum X_1^2 & \sum X_1X_k \\ \sum X_k & \sum X_kX_1 & \sum X_k^2 \end{pmatrix}$$

The variance covariance matrix of the least squares estimates are given by

$$(X^1X)^{-1}\, s_e^2 = \begin{pmatrix} var(\beta_0) & cov(\beta_0\beta_1) & cov(\beta_0\beta_k) \\ cov(\beta_1\beta_0) & var(\beta_1) & cov(\beta_1\beta_k) \\ cov(\beta_k\beta_0) & cov(\beta_k\beta_1) & var(\beta_k) \end{pmatrix}$$

The S.e. of β_j is the square root of the j^{th}diagonal element of $(X^1X)^{-1}\, s_e^2$

Confidence Interval:

The confidence interval for individual β_j is given by

$\boldsymbol{\beta_j}$ - $t_{\frac{\alpha}{2}}$Se $\beta_j \leq \boldsymbol{\beta_j} \leq \boldsymbol{\beta_j} + t_{\frac{\alpha}{2}}$Se $\boldsymbol{\beta}$**j**

Rejection region:

We reject H_0 if $-t_c < -t_{\frac{\alpha}{2}}$ or $t_c > t_{\frac{\alpha}{2}}$ or if Pvalue$\ll\propto$

CHAPTER FOUR: DATA ANALYSIS AND RESULTS

4.1Descriptive Analysis

In my literature review, I have classified violent crimes into two i.e.

(i) Crimes against Person and

(ii) Crimes related to property.

I found the total number of violent crimes by summing up the two types of violent crimes per year. Table 4.1a shows the estimated values of annual Population. Table 4.1b shows the computed values of crime rate and the seven independent variables for 15 years (i.e. n=15).

Table 4.1a: Annual Population Estimates

Year	1999	2000	2001	2002	2003	2004	2005
Population	28700000	29563268	30452502	31368483	32312016	33283930	34285078

2006	2007	2008	2009	2010	2011	2012	2013
35316339	36378620	37472853	38600000	39761050	40957023	42188970	43457973

Table 4.1b: Crime rate and Independent variables

y	x_1	x_2	x_3	x_4	x_5	x_6	x_7
135.8885	2.525289	91.806	13.69652	2.17E+09	0.020599	37.62021	19.13798
144.2297	2.564078	72.974	13.46047	2.27E+09	0.018517	40.90541	19.97162
135.2401	2.606242	76.077	13.30196	2.49E+09	0.01732	38.40196	20.80686
122.0748	2.592491	70.229	13.22057	2.7E+09	0.014136	45.12025	21.68228
60.85662	2.646954	70.056	13.18865	3.01E+09	0.014546	53.57362	22.51074
143.3815	2.756396	98.336	10.35685	3.73E+09	0.017146	54.25	23.80506
122.7852	2.692072	58.570	10.62392	4.03E+09	0.018317	45.67435	24.51009
115.7	2.877318	88.268	10.91397	4.49E+09	0.01626	47.07263	25.12123
99.05268	3.189911	75.640	11.08027	4.91E+09	0.014597	40.01081	25.61865
94.28692	3.618353	64.414	11.79503	5.49E+09	0.011537	42.27749	26.03717
106.5466	3.737563	64.805	10.47005	6.3E+09	0.014063	46.12437	27.16523
100.5331	4.062408	90.568	10.00197	6.68E+09	0.012484	37.0344	28.15012
102.82	4.20881	105.400	9.456905	7.71E+09	0.011069	33.2381	28.8481
102.2661	4.422171	61.142	9.835104	8.37E+09	0.009656	30.00231	29.51963
93.29703	4.707606	75.595	9.428412	9.09E+09	0.009655	34.00224	30.25682

Where:

y = Annual Crime rate per 100,000

x_1 = percentage population enrolled for Secondary Education

x_2 = Number of Fire Arms $ Ammunitions per 100,000

x_3 = Number of Police officers per 100,000

x_4 = Gross Domestic Product per 100,000

x_5= Percentage of population in possession of dangerous drugs

x_6 = Percentage of population convicted of violent crimes

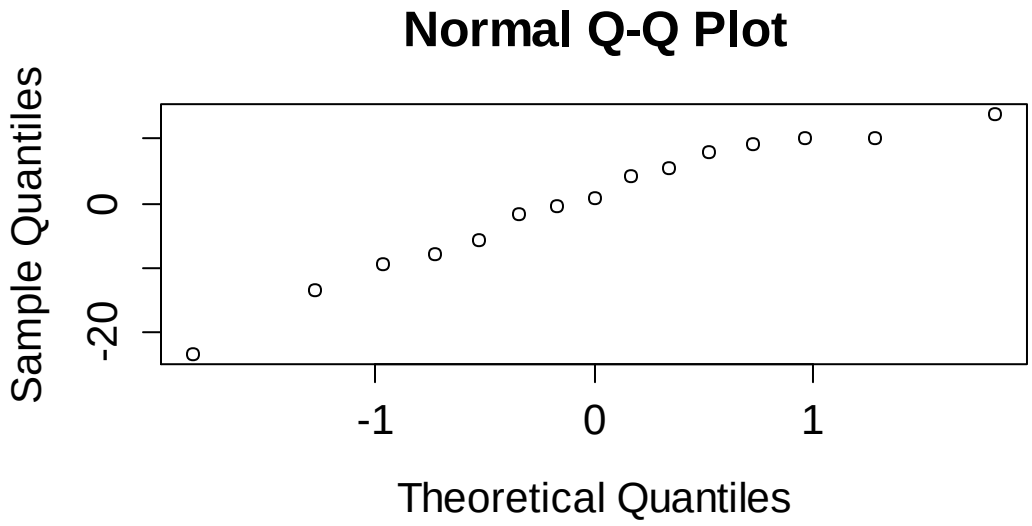
x_7 = Percentage of population employed

4.2 Multiple Linear Regressions:

I subjected these variables to Multiple linear regression using R statistical software and the output is as shown in figure 4.2 and table 4.2 both below.

4.2.1 Model I:

Figure 4.2: Normal Q-Q Plot1



The result of the Q-Q Plot indicates that the residuals came from a normal distribution. This is proof that the data belongs to a normally distributed function.

Table 4.2: Multiple linear regression Out put1

Min. 1st Qu. Median Mean 3rd Qu. Max.							
Call:							
lm(formula = $Y \sim X_1 + X_2 + X_3 + X_4 + X_5 + X_6 + X_7$)							
Residuals:							
Min 1Q Median 3Q Max							
-23.2279 -6.7406 0.7823 8.5213 13.8562							
Coefficients:							
	Estimate	Std. Error	t value	Pr(>|t|)			
(Intercept)	1.055e+03	4.554e+02	2.317	0.0536 .			
X_1	1.758e+01	4.494e+01	0.391	0.7073			
X_2	-8.233e-04	5.437e-02	-0.015	0.9883			
X_3	-3.282e+01	1.276e+01	-2.573	0.0369 *			
X_4	-3.365e-09	2.175e-08	-0.155	0.8814			
X_5	-2.250e+03	4.480e+03	-0.502	0.6310			
X_6	-1.201e+00	1.197e+00	-1.003	0.3493			
X_7	-2.121e+01	1.094e+01	-1.938	0.0938 .			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1							
Residual standard error: 14.61 on 7 degrees of freedom							
Multiple R-squared: 0.7912, Adjusted R-squared: 0.5824							
F-statistic: 3.789 on 7 and 7 DF, p-value: 0.04994							

4.2.2Significance of the model fit:

With a P value of 0.04994 and R^2 =79%, the model fit is statistically significant. We reject the null hypothesis; H0: all $\beta_j = 0$, and conclude that there is no sufficient evidence to conclude that the predictor variable X_j adds zero predictive value to the model fitted. ; j = 0, 1, 2, 3, --------

The model below resulted from the output above:

$$Y = 1055 + 17.58X_1 - 0.0008233X_2 - 32.82X_3 - 3.365 \times 10^{-9}X_4 - 2250X_5 - 1.201X_6 - 21.21X_7$$

Where:

$y, x_1, x_2, x_3, x_4, x_5, x_6, x_7$ are as on page 18.

The model shows that only one out of the seven variables fitted was statistically significant. This is the Police ratio. Employment rate is statistically significant at 90% confidence level i.e.$\alpha = 0.1$.

4.2.3 Interpreting model Parameters

95% Confidence intervals:

Using values of $t_{\frac{\alpha}{2}}$ and s.e obtained from the multiple linear regression output1, the confidence intervals were determined as follows.

β_0 : $0 \leq \beta_0 \leq 2110$

β_1: $0.01 \leq \beta_1 \leq 35.15$

β_2 : $-7.75 \times 10^{-9} \leq \beta_2 \leq -1.6 \times 10^{-3}$

β_3 : $-0.01 \leq \beta_3 \leq 65.65$

β_4: $-6.736 \times 10^{-9} \leq \beta_4 \leq 6 \times 10^{-12}$

β_5: $-4499 \leq \beta_5 \leq -1.1$

β_6 : $-2.4 \leq \beta_6 \leq -4.09 \times 10^{-4}$

β_7: $-42.41 \leq \beta_7 \leq -0.008$

From this model:

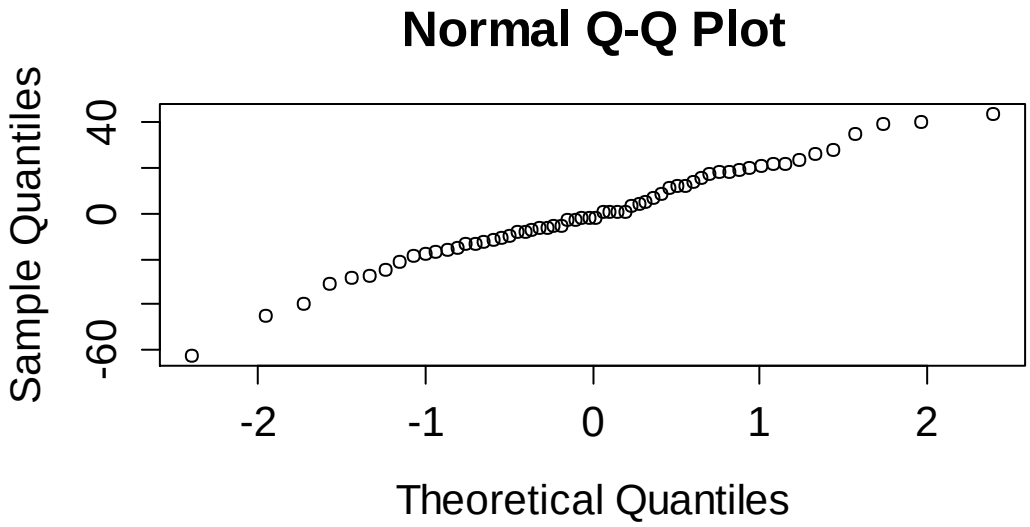
- The mean annual crime rate when all the predictor variables have value = 0 is 1055 crimes per 100,000.
- Increasing secondary school enrolment by 1% increases the Crime rate by 17.58crimes per 100,000 per year when all other variables in the model are kept constant.
- Increasing the number of fire arms and ammunitions by one unit decreases the violent crime rate by 0.0008233 crimes per 100,000 per year while keeping all other variables constant.
- An increase in Police ratio of 1 unit results in a decrease in violent Crime rate of 32.82 crimes per 100,000 of the population per year while keeping all other variables constant.
- An increase in Gross domestic product of ksh1 per 100,000 of population decreases the rate of violent crime by 3.365×10^{-9}crimes per 100,000 of population per year while keeping all other variables constant.
- When the percentage of people found in possession of dangerous drugs increases by 1%, the rate of violent crime decreases by 2250 crimes per 100,000 of the population per year while keeping all other variables constant.
- Increasing conviction rate by 1% results in a decrease of 1.2 violent Crimes per 100,000 of the population per year while maintaining all other variables in the model constant
- Raising the employment rate by 1% decreases the rate of violent crime by 21.2 violent crimes per 100,000 of the population per year while holding all other variables in the model constant.

4.2.4 Model II:

Due to insufficient amount of data, I have simulated more data to increase the sample size from 15 to 60.

The results of the output are shown in figure4.3 and table 4.3 below.

Figure 4.3: Q-Q Plot2



The result of the Q-Q Plot indicates that the residuals came from a normal distribution. This is proof that the data belongs to a normally distributed function.

Table 4.3: Multiple linear regression output 2:

```
Call:
lm(formula = Rate ~ SSC + FIR + POL + GP + DAB + CNV + EPR)

Residuals:
    Min      1Q  Median      3Q     Max
-36.030 -14.588   0.423  14.331  41.272

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.374e+02  3.611e+01   3.806 0.000374 ***
SSC          2.387e+00  3.601e+00   0.663 0.510375
FIR          5.006e-01  1.687e-01   2.967 0.004537 **
POL         -1.375e+00  1.589e+00  -0.865 0.391005
GP          -2.595e-09  1.344e-09  -1.931 0.058928 .
DAB         -1.302e+03  6.899e+02  -1.888 0.064663 .
CNV         -8.048e-01  3.796e-01  -2.120 0.038808 *
EPR          1.750e-01  6.964e-01   0.251 0.802606

Residual standard error: 19.63 on 52 degrees of freedom
```

Table 4.3 shows that:

(i) The median = 0.423($very\ small$) .This implies that mean (ε_i) ≅ 0.Hence the condition for linearity is satisfied.

(ii) Out of the seven predictors variables in the model, only two are statistically significant to the model fit at 95% level of confidence ($\propto = 0.05$).These are :
Fire arms and ammunitions and the percentage of the population convicted of violent crimes.

(iii) Two other predictor variables i.e.GDP and the percentage of the Population found in possession of dangerous drugs are statistically significant at lower level of confidence i.e.90 %($\propto = 0.1$).
The model below resulted from the output above.

$$Y = 137.4 + 2.387X_1 + 0.5X_2 - 1.375\ X_3 - 2.595 \times 10^{-9}X_4 - 1302X_5 - 0.804X_6 + 0.175X_7$$

Where: $y, x_1, x_2, x_3, x_4, x_5, x_6, x_7$ are as on page 18.

4.2.5Significance of the model fit

Results of the output show that the overall model test gives a P value of 0.03511 hence

P value $\ll$ 0.05. This means that we reject the null hypothesis H_0: all β_{js} = 0

R^2 = 24%. Means that when sample size, n=60 then only 24% of variability is explained by the model.

4.2.6 Interpreting model Parameters

95% Confidence intervals:

Using values of $t_{\frac{\alpha}{2}}$ and s.e obtained from the multiple linear regression output2, the confidence intervals were determined as follows.

β_0 : -0.03466≤ $\beta_0 \leq$ 274.8347

$\beta_{1:}$ -0.00046 ≤ β_1 ≤ 4.774463

β_2 : -0.00053≤ β_2 ≤ 1.00053

β_3 : -2.74949 ≤ β_3 ≤ -0.00051

β_4: -5.19026× $10^{-9} \leq$ β_4 ≤ 0.00026 × 10^{-9}

β_5: $-2604 \leq \beta_5 \leq 0.$

β_6: $-1.60875 \leq \beta_6 \leq 0.00075$

β_7: $0.000204 \leq \beta_7 \leq 0.349796$

From this model:

- The mean annual crime rate when all the predictor variables have value = 0 is 137.4 crimes per 100,000
- Increasing enrolment for secondary education by 1% increases the Crime rate by 2.387 crimes per 100,000 per year when all other variables in the model are kept constant.
- Increasing the number of fire arms and ammunitions by one decreases the violent crime rate by 0.0008233 crimes per 100,000 per year while keeping all other variables constant.
- An increase in Police ratio of 1 unit results in a decrease in violent Crime rate of 32.82 crimes per 100,000 of the population per year while keeping all other variables constant.
- An increase in Gross domestic product of ksh1 per 100,000 of population decreases the rate of violent crime by 3.365×10^{-9}crimes per 100,000 of population per year while keeping all other variables constant.
- When the percentage of people found in possession of dangerous drugs increases by 1%, the rate of violent crime decreases by 2250 crimes per 100,000 of the population per year while keeping all other variables constant.
- Increasing conviction rate by 1% results in a decrease of 1.2 violent Crimes per 100,000 of the population per year while maintaining all other variables in the model constant
- Raising the employment rate by 1% decreases the rate of violent crime by 21.2 violent crimes per 100,000 of the population per year while holding all other variables in the model constant.

4.2.7 MODEL III: Crime Rate Versus Significant Predictors:

This is a model of the response variable regressed on the two significant predictors (i.e fire arms and ammunitions + conviction rate)

The output is shown in table 4.4below:

Table 4.4 Multiple regression output(Crime rate Vs Significant Predictors)

```
Call:
lm(formula = Rate ~ FIR + CNV)

Residuals:
    Min      1Q  Median      3Q     Max
-65.987 -13.921   1.288  13.625  57.672

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 122.5638    21.2166   5.777 3.33e-07 ***
FIR          -0.3473     0.1727  -2.011    0.049 *
CNV           0.1871     0.4125   0.454    0.652
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 22 on 57 degrees of freedom
Multiple R-squared:  0.06762,   Adjusted R-squared:  0.03491
F-statistic: 2.067 on 2 and 57 DF,  p-value: 0.1359
```

Table 4.4 shows that:

(i) Out of the two predictors variable in the model, one is statistically significant to the model fit at 95% level of confidence($\propto = 0.05$).This is

Fire Arms and Ammunitions. However the P value =0.1359 hence> 0.05 .We do not reject the null hypothesis H_0: β_{js} = 0.The model is not statistically significant. The low value of R^2 also does not make sense.

The model below resulted from the output above.

$$y = 122.5638 + 0.1871\, X_6 - 0.3473 X_2$$

From this model;

(i) The mean annual rate of violent crime when both predictor variables have value zero is 122.5638 crimes per 100,000.

(ii) Increasing the conviction rate by 1% increases the crime rate by 0.1871 violent crimes per 100,000 while holding the number of fire arms and ammunitions constant.

(iii) If the number of fire arms and ammunitions recovered or surrendered increases by 1 then the rate of violent crime will decrease by 0.3473 crimes per 100,000 per year.

4.2.8 Discussion

The results obtained using model I show that there is only one variable out of the seven predictor variables which is statistically significant. This implies that using this model, there is only one key factor of violent crime rate. This is the Police ratio:

Oliver'02, Alison (2002)"The Economics of crime: An analysis of crime rates in America," also found that Police ratio was statistically significant as a factor of crime rate. These results agree to a greater extent with the literature review.

The results obtained using model II show that two variables , number of fire arms and ammunition and also the conviction rate are statistically significant.

Jim Frost (12June,2014),"How to interpret a Regression model with Low R-squared and Low P values" says that: (a) the interpretation of the significant variables are the same for both high and low R-squared models.(b) Even when R-squared is low, low P values still indicate a real relationship between the significant predictors and the response variable.

However, some of the results obtained using both model I and model II are not good reflections of the literature review. For example, in both models the following predictor variables do not agree with the literature review.

(i) The percentage of enrolment for secondary education and

(ii) The percentage of people found in possession of dangerous drugs.

CHAPTER FIVE: CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

This research study has found out that the higher the Police ratio, the lower the rate of violent crime. In the model above, when you increase the Police ratio by 1 per 100,000 of the Population, the rate of violent crime decreases by 32.82 crimes per 100,000 per year keeping other variables constant.

This study has also established that increasing the conviction rate by 1% increases the crime rate by 0.1871 violent crimes per 100,000 per year while holding the number of fire arms and ammunitions constant. However this result does not agree with the literature review since conviction rate was expected to be a deterring variable.

If the number of fire arms and ammunitions recovered or surrendered increases by 1 then the rate of violent crime will decrease by 0.3473 crimes per 100,000 per year while keeping conviction rate constant.

5.2: Recommendations

This study therefore suggests that:

(i) There is need to invest heavily in recruiting more Police officers to increase the ratio of number of Police officer per 100,000 of the population.

(ii) There is great need to keep conviction rates low in order to keep violent crime rates low.

(iii) There is need to put more emphasis in recovery and surrender of fire arms and ammunitions so as to reduce violent crime rates in the country.

(iv) Finally, there is still a lot of room for research in this area of study. I wish to encourage more scientists to take keen interest in studying this area of research.

REFERENCES

Arrigo, Bruce A., ed. *Social Justice, Criminal Justice.* Belmont, CA: Wadsworth, 1999.

Belmont, CA: Wadsworth Thomas Learning, 2001.

Bowlby, John (1988). *A Secure Base: Parent-Child Attachment and Healthy Human Development.* New York: Basic Books.

Cleckely, Hervey (1982) . *The Mask of Sanity.* New York: New American Library.

Cohen, Albert K (1955). *Delinquent Boys: The Culture of the Gang.* New York: Free Press.

Curran, Daniel J., and Claire M. Renzetti. *Theories of Crime.* Boston, MA: Allyn & Bacon, 2001.

Ehrlich, Isaac.(1973) .Participation in Illegitimate Activities: A Theoretical and Empirical Investigation.. The Journal of Political Economy, Vol. 81.

Ernst Freund 1915, Classification and Definition of crimes; Journal of criminal law and criminology.

Fender, John (1999). .A General Equilibrium Model of Crime and Punishment.. Journal of Economic Behavior & Organization. Vol. 39.

Fleisher, Mark S.(1995) *Beggars and Thieves: Lives of Urban Street Criminals.* Madison, WI: University of Wisconsin Press.

Gimode 2001,An A anatomy of violent crime and Insecurity in Kenya: The case study of Nairobi 1985-1999.

Jost, Peter J. (2001) Crime, Coordination, and Punishment: An EconomicAnalysis.International Review of Law & Economics.Vol. 21.

Journal of Economic Perspectives. Vol. 10, Winter 1996, 43-67.

Journal of Political Economy 76, Mar.-April 1968), 169-217. Davis, Michael LTime and Punishment: An Intertemporal Model of Crime.. The Journal of Political Economy, 96, April 19988, 383-390. Ehrlich, Isaac .Crime, Punishment, and the Market for Offenses.

Karr-Morse, Robin, and Meredith S. Wiley (1997). *Ghosts from the Nursery: Tracing the Roots of Violence.* New York: Atlantic Monthly Pres

Levitt, Steven D (1998) .Why do Increased Arrest Rates Appear to Reduce Crime: Deterrence, Incapacitation, or Measurement Error? Economic Inquiry. Vol. 36.

Lombroso, Cesare. (1968) *Crime: Its Causes and Remedies.* Montclair, NJ: Patterson Smith.

Martinez-Vazquez, Jorge; Seaman, Bruce A. (1997) Criminal Behavior in General Equilibrium: Who Benefits from Crime?. Journal of Socio-economic Vol. 26. Neilson, William S; winter, Harold. .On Criminals. Risk Attitudes.. Economics Letters. Vol. 55, August 1997, 97-102.

Oliver'02 Alison (200); The Economics of crime: An Analysis of crime rates in America.

Ramanathan, Ramu. (1998) Introductory Econometrics with Applications. San Diego: Harcourt Brace College Publishers.

Renzetti, Claire M., and Lynne Goodstein, eds (2001). *Women, Crime, and Criminal Justice.* Los Angeles: Roxbury. Becker, Gary S. Crime and Punishment: An Economic Approach.

Siegel, Larry. Criminology: Theories, Patterns, and Typologies.