

Modelling Crime Rate Using a Mixed Effects Regression Model

Chris Muchwanju^{*}, Joel Cheruyiot Chelule, Joseph Mung'atu

Department of Statistics and Actuarial Science, Jomo Kenyatta University of Agriculture and Technology (JKUAT), Nairobi, Kenya

Email address:

chrismuchwanju@gmail.com (C. Muchwanju)

To cite this article:

Chris Muchwanju, Joel Cheruyiot Chelule, Joseph Mung'atu. Analysis of Crime Rate Using a Mixed Effects Regression Model. *American Journal of Theoretical and Applied Statistics*. Vol. 4, No. 6, 2015, pp. 496-503. doi: 10.11648/j.ajtas.20150406.20

Abstract: In this paper we propose a type of Mixed effects Regression Model, that is Hierarchical Linear Model to study crime rate. We derive the estimators of the proposed model and discuss the asymptotic properties of the model. In order to test for the practicability of the proposed model, we estimate a crime equation using a panel dataset of the provinces in Kenya for the period 1992 to 2012 employing the REML estimator. Our empirical results suggest that Poverty Rate, Unemployment rate, Probability of arrest, population Density and police rate are correlated to all typologies of crime rate considered. The results further suggest that crime rate is better explained at provincial level as compared to country level.

Keywords: Mixed Effects Model, Panel Data, Crime Rate, Kenya, Provinces

1. Introduction

Crime is one of the major problems of any society. It causes disequilibrium in the social system. It is a complex phenomenon of social, political and economic impacts. More important, crime can be seen as an indirect symptom of failures in personal development, socialization and education of children- failures which represent very large losses of human potential.

The reasons why people are more or less law-abiding are manifold. Crime is as old as mankind itself, moreover, it has become a common societal phenomenon and many people view it as if it is a functional component of the organization of human groupings, Schafer (1976).

Despite the large research literature addressing various aspects of these issues, there remains limited evidence giving both a broad picture of the range of anti-social activity and a detailed longitudinal picture of the development processes leading to criminal outcomes.

1.1. Related Works

Since Becker (1968) published his seminal work on the economics of crime, a large empirical literature has been developed around estimation and testing of economic model of crime. Almost all of these contributions have used aggregate data, usually at state or national level. Economists have devoted considerable effort to determining

its empirical validity.

Ideally, the economic model of crime should be estimated with individual level data since the model purports to describe the behavior of individuals. However, the expense and difficulty to create random sample which is large enough to include representative information about individual criminal activity has been, and still is a challenge. This challenge has continued to be obstacle for individual level analysis. Few literatures have used individual data, but empirical result failed to obtain the expected estimate, Christopher and Trumbull (1994).

In the absence of empirical work at the individual level, the estimation of the economic model of crime using aggregate data continues Craig (1987), Avio (1988) and Trumbull (1989). While estimation with aggregate data has been criticized, results from such estimation have influenced public policy.

The consensus of the empirical literature is that punishment (certainty and severity), Incarceration population, police rate and unequal opportunities have strong deterrents effect.

Estimates of the magnitude of deterrence effect vary, but it appears that an increase in law enforcement activity that increase the probability of punishment or the severity of punishment. Further empirical investigation is necessary in order to gain a more accurate estimate of the magnitude of this deterrent effect.

Objections to economic studies of criminal behavior have been many and occasionally fierce; see for example, Blumstein,

Cohen and Nagin (1978), Orsagh (1979), Brier and Fienberg (1980), Prisching (1982), and Cameron (1988). In particular, studies based on aggregate data have been criticized. In addition to attacks on the assumption of rational behavior, the main criticism relates to interpretation of empirical results, method of estimation, statistical identification of equations, and unobserved heterogeneity, measurement errors, and operationalization of theoretical variables. The effects of various economic factors are less clear.

The vast majority of the research has been theoretical in emphasis e.g. Stigler (1970), Posner (1977), Andreoni (1991), Kaplow and Shavell (1999). Less progress has been made in empirically testing the economic model of crime, although there are a number of notable attempt to do so (e.g. Ehrlich (1973), Witte (1980), Cameron 1988, Tauchen, Witte and Griensinger 1994). One major difficulty in testing the Becker (1968) model and its numerous extensions is that many of the predictions of the model are empirically indistinguishable from other competing models. For example, except under special circumstances, it is difficult to separate deterrence (the basis of the economic model).

Many supposed tests of the economic model of crime have little power to discriminate between competing models. Difficulty that arises in testing the economic model of crime is identifying exogenous sources of variation in the criminal justice system that is necessary to identify a causal link between response and deterrent variables. The early empirical literature in this area, lacking exogenous variation, yield results that are difficult to interpret and have been harshly criticized Fisher and Nagin (1978), Cameron (1988).

Cornwell and Trumbull (1994) point to the fact that aggregate cross-section econometric techniques do not control for unobserved heterogeneity. Addressing this problem by use of a panel data dataset of North Carolina counties, they obtain more modest deterrent effects of the arrest and conviction rates than those obtained from cross-section estimation.

Crime is among the most difficult of the many challenges facing Kenya. The country's crime rates are among the highest in the world and no Kenyan is insulated from its effects. Beyond the pain and loss suffered by crime victims, crime also has less direct costs. The threat of crime diverts resources to protection efforts, exacts health costs through increased stress, and generally creates an environment that is not conducive to productive activities relationship with returns from crime and non-crime activities. All of these effects are likely to discourage investment and stifle long-term growth in Kenya.

There is huge potential to fill in such gaps in knowledge based quantitative research from the very rich and largely unexploited register data that exists in the African countries, in Kenya in particular.

Despite of its remarkable features, Kenya criminal activity has received little attention and remains largely neglected by the economies of crime of literature. Indeed, if one goes through this copious literature one will hardly find an allusion to Kenya and other African countries.

In this proposal, we present no empirical evidence has been attempted to deter crime in Kenya. Thus, the ability of the criminal justice system to deter crime is very weak in the past long year. Hence we use the available panel data of crime for the provinces in Kenya for the period 1992 to 2012 to model crime rate in Kenya. The result of this finding will help policy makers and planners to target with feasible interventions and strategies to reduce crime in Kenya.

1.2. Statement of the Problem

Crime is on the rise globally, but is more accelerating in African countries. In particular, Kenya ranked the second-highest after South Africa on the survey on economic crimes in the world (UN, 2009). Criminal behavior is steadily increasing and over half of the population worries about crime constantly and roughly 75% feel unsafe while at home (UN-Habitat, 2002).

Many factors have been attributed to the ever increasing crime rate in Kenya. Crime statistics in Kenya over the years has remained at the level of summary, tabulation and lacks scientific analysis that allows for prediction, planning and effective control. There is a need for a policy change towards crime reduction in the country. Statistical modeling can be very useful policy making in the management of crime in Kenya. In this research project we are going to consider provincial crime rate data for crimes against person and property in Kenya.

1.3. Objectives

1.3.1. General Objective

To model crime rate in Kenya using a mixed effect regression model

1.3.2. Specific Objectives

- i. To propose a mixed effects regression model.
- ii. To estimate the proposed model using a panel dataset of Kenyan provinces for the period 1992-2012.

1.4. Strengths and Weaknesses of the Research

The research has a number of strengths. First, data was obtained from different organizations and was believed to be reliable. This design provided a clearer understanding on crime rate in the study area which was the entire country Kenya. During the analysis there was no missing data and as such all the variables were analyzed without being affected by missing data. However, the study had some weaknesses, since we used secondary data, there was a potential for missing some vital information on crime due to under reporting of crime in some areas.

2. Review of the Previous Studies

2.1. Introduction

Crime is a social problem which has direct effects, generating large losses to victims. For instance, loss of property, damage of property, injuries to victims and stress-

related complications; and to the government through the cost of maintaining the criminal justice system, policing, correction and health systems, Entorf and Spengler (2002).

2.2. Crime Rate

Galloway and Pudney (2011) saw crime as an indirect symptom of failures in personal development, socialization and education of children-failures which represent very large losses of human potential. They further observe that despite the large research literature addressing various aspects of this problem, there still remains very limited evidence giving both a broad picture of the range of anti-social activity and a detailed longitudinal picture of the development processes leading to criminal outcomes.

More recently, some progress has been made in this area. For instance, on the question of whether more police reduce crime, a series of papers using a range of different approaches have all come to a similar conclusion that more police substantially reduce crime, Marvell and Moody (1968), Levitt (1977) and Corman and Mocan (2000).

During the last 50 years economists have devoted considerable effort to determining its empirical validity since Becker's (1968) have published his seminal article on the economic model of crime and they have invaded the field using their all-embracing model of individual rational behavior, where a criminal act is preferred and chosen if the total pay-off, including that of sanctions and other costs, is higher than that of legal alternatives. Much of this research examines deterrence, the idea that policy can reduce crime by raising the expected costs, Erling (1999).

The study of crime has always been a multidisciplinary activity. Along history different school of thought have proposed different and sometimes convictive ways of considering time. Apart from criminologists, sociologists are perhaps the dominant group, but psychologists and political scientists have also long been prominent. Economists, econometricians and statistician are among the most recent entrants. With Becker's 1968 "Crime and Punishment: An Economic Approach," serving as the starting point for modern economists' work on crime and Cornwell and N. Trumbul (1994) and Marselli (1997) "Estimating the Economic Model of Crime with panel data." In the mid to late 1990's, there are renewed flurry of work by young economists who developed research agendas largely centered on the study of crime.

Crime is among the most difficult of the many challenges facing Kenya. The country's crime rates are among the highest in the world and no Kenyan is insulated from its effects (UN, 2002, 2009, 2011). Beyond the pain and loss suffered by crime victims, crime also has less direct costs. The threat of crime diverts resources to protection efforts, exacts health costs through increased stress, and generally creates an environment unconducive to productive activity.

2.3. Research Gap

Crime threatens human security and rights, undermines

economic, political and social development of various countries worldwide. Despite vast budgetary allocations by the government to fight crime, little has been achieved so far and this has impacted the country's development agenda negatively. According to official crime data provided, there has been a drastic rise in crime in the last three decades in most developing countries, Kenya included.

Indeed, Kenya has seen a tremendous growth in social, economic and political field, but unless there are clear policy guidelines to deal with issues of crime, the vision 2030 may not be realized. Crime is a pertinent issue in Kenya and has persisted thus needs a deeper study.

The heterogeneous report from the police and other law enforcement agencies leaves doubt on why such trends thrive. It is therefore, crucial to study crime rate and the reason why it increases, not only for academic purpose but also for designing and diligently prescribing effective policies for planning, predicting and preventing crime.

Despite its remarkable features, Kenya's criminal activity has received little attention and remains largely neglected by the economics of crime literature. There is huge potential to fill in such gaps in knowledge based quantitative research from the very rich and largely unexploited registered data that exists in African countries, Kenya in particular. Indeed, if one goes through this copious literature one will hardly find any allusion to Kenya or to any other African countries.

3. Methods

3.1. Study Area and Data Source

Kenya is Located in Africa, East Africa, and it is divided into eight administrative units called provinces. These provinces are of different sizes and also vary in the population of people who stay there.

The data was obtained from annual Statistical abstract and reports from KNBS, Kenya Police annual report and the World Bank Reports.

3.2. Study Variables

The dependent variable is the crime rate which is defined as the crimes reported per 100,000 population. Independent Variables are poverty rate, unemployment rate, school enrolment rate, percentage young men (15-29 years), conviction rate, clear up rate, probability of arrest, population density and police rate.

3.3. Statistical Analysis Methods

3.3.1. Introduction

Crime rate is a function of many factors or variables called explanatory or independent or predictor. Thus, crime rate is the performance measure called the response or dependent variable. Hence, the relationship between crime rate and explanatory variables takes the form of a multiple regression model.

3.3.2. Review of Multiple Regression

In general, a multiple regression model takes the form;

$$Y = f(\xi_1, \xi_2, \dots, \xi_k) + \varepsilon \quad (1)$$

Where Y is the response variable, $\xi_1, \xi_2, \dots, \xi_k$ are the independent variables and ε is a term that represents other sources of variability not accounted for in the function f . This ε may include effects such as measurement errors on the response, background noises and even effects of other variables. It is treated as statistical error $\sim N(0, \sigma^2)$.

Consequently;

$$E[Y | \xi_1, \xi_2, \dots, \xi_k] = E[f(\xi_1, \xi_2, \dots, \xi_k)] + E[\xi_1, \xi_2, \dots, \xi_k] = E[f(\xi_1, \xi_2, \dots, \xi_k)] \quad (2)$$

The variables $\xi_1, \xi_2, \dots, \xi_k$ in equation (3.1) are called natural variables because they are expressed in the natural units in which the measurements being studied were made. It is convenient to transform these natural variables into coded variables, say X_1, X_2, \dots, X_k , which are dimensionless with mean zero and some standard deviation. Accordingly, in terms of the coded variables, the response function (1) can be written as;

$$Y = f(x_1, x_2, \dots, x_k) + \varepsilon \quad (3)$$

where ε are random variables called error terms which are assumed to be identically and independently distributed, independent of X and normally distributed with zero mathematical expectation i.e. $E(\varepsilon) = 0$, and constant and finite variance i.e. $Var(\varepsilon) = \sigma^2 < \infty$. The explanatory variables X are assumed to be non-random, Hosmer (2013)

Since the true response function f is unknown, it is approximated. In its approximation, the efficiency of the estimation procedure depends on the ability to develop a suitable approximation for this function. This tenability of an efficient approximation is usually the focus in modeling.

3.4. Mixed Effect Model

In light of the explanation provided in sub-section 3.3.2, together with the reason that panel data may have group effects, time effects, or both, whereby these effects are either fixed or random, we propose a Mixed-Effects Regression Model that takes into account the two levels of analysis, that is at provincial level and at country level. We therefore propose a Hierarchical Linear Model of the form;

i). One at provincial level

$$\text{Level 1: } Y_{ij} = \beta_{0j} + \beta_{1j}X'_{ij} + e_{ij} \quad (4)$$

ii). At country level

Level 2:

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + \mu_{0j} \\ \beta_{1j} &= \gamma_{10} + \mu_{1j} \end{aligned} \quad (5)$$

Where;

Y_{ij} is the crime rate, X'_{ij} contains a set of control variables

and deterrent variables, and μ_{0j}, μ_{1j} and ε_{it} are error terms assumed to be iid with zero expectation and $Var(\mu_{0j}) = \sigma_{\mu}^2$, $Var(\mu_{1j}) = \sigma_{\mu}^2$ and $Var(\varepsilon_{it}) = \sigma_{\varepsilon}^2$

In this mixed effects model, the effect of the independent variables are allowed to have varying effects across level 2 units because their coefficients in level 2 of the model includes the random error terms. It is these error variances that allow the effect of the independent variables to take on different values across level 2.

The equation for the intercept β_{0j} consists of the overall mean intercept γ_{00} and a cluster-specific random intercept only μ_{0j} . The additional equations for the slopes β_{1j} consists of the overall mean slopes γ_{10} and cluster specific random slopes μ_{1j}

Thus combining (4) and (5) above gives the overall mixed effects model;

$$\begin{aligned} Y_{ij} &= \gamma_{00} + \mu_{0j} + (X'_{ij})(\gamma_{10} + \mu_{1j}) + e_{ij} \\ Y_{ij} &= \gamma_{00} + \gamma_{10}X'_{ij} + \mu_{0j} + \mu_{1j}X'_{ij} + e_{ij} \end{aligned} \quad (6)$$

3.5. Estimation of HLM

We now describe the estimation techniques for our two-level hierarchical model. We now present notation which we use throughout this section. Consider the level 1, level 2 and combined models (using HLM) shown in (4), (5) and (6), respectively. Using matrix notation these models are represented as follows:

$$\text{Level 1: } Y_i = \beta_j X'_i + e_{ij}, j = 1, 2, \dots, J \quad (7)$$

Where;

$$Y_j = \begin{bmatrix} Y_{1j} \\ Y_{2j} \\ Y_{nj} \end{bmatrix}, X_j = \begin{bmatrix} 1 & X_{1j} \\ 1 & X_{2j} \\ 1 & X_{3j} \end{bmatrix}, \beta_j = \begin{bmatrix} \beta_{0j} \\ \beta_{1j} \end{bmatrix}, \varepsilon_j = \begin{bmatrix} \varepsilon_{1j} \\ \varepsilon_{2j} \\ \varepsilon_{nj} \end{bmatrix}$$

$$\text{Level 2: } \beta_j = \gamma + \mu_j \quad (8)$$

Where;

$$\beta_j = \begin{bmatrix} \beta_{0j} \\ \beta_{1j} \end{bmatrix}, \gamma = \begin{bmatrix} 1 & \gamma_{00} \\ 1 & \gamma_{10} \\ 1 & \gamma_{11} \end{bmatrix}, \mu_j = \begin{bmatrix} \mu_{0j} \\ \mu_{1j} \end{bmatrix}$$

Combined model;

$$Y_{ij} = \gamma X'_j + \mu_j X'_j + \varepsilon_{ij} \quad (9)$$

X is a design matrix, γ is a vector of fixed effects, μ_j is a vector of random effects and ε_{ij} is a vector of random errors. The assumptions are as outlined below;

$$\varepsilon_{ij} \sim N(0, R), R_j = \delta^2 I_{n_j}$$

$$\mu_j \sim N(0, G), G = \begin{bmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{bmatrix} \quad (10)$$

Several estimation techniques are used in hierarchical linear modeling since the model comprises different types of parameters. Specifically, the level 1 coefficient, β_j can be

fixed, that is equal to a constant β_{1j} , non-randomly varying (that is across level 2).

The level 2 coefficients, γ , are considered fixed effects and the level 1 and level 2 variances and covariances (δ^2 , τ_{00} , τ_{11} and τ_{11}) are called the covariance components.

The estimation techniques for each type of parameter are outlined below.

3.5.1. Estimating Fixed Effects (γ)

Weighted least squares (WLS) or generalized least squares (GLS) is used to estimate γ as shown below:

$$\hat{\gamma} = (A^T \hat{V}^{-1} A)^{-1} A^T \hat{V}^{-1} Y \quad (11)$$

$$V = \text{Var}(Y) = XGX^T + R \quad (12)$$

A is a $N \times 4$ design matrix with $= \sum_{j=1}^J n_j$, and \hat{V} is V with G and R replaced by their Maximum Likelihood Estimates. The elements of G and R (that is δ^2 , τ_{00} , τ_{11} and τ_{11}) are called the variance covariance components and are estimated by Maximum Likelihood or Restricted (ML) or restricted maximum likelihood (REML) as described below

The variance of the estimator $\hat{\gamma}$ in (10) is estimated by

$$\widehat{\text{Var}}(\hat{\gamma}) = (A^T \hat{V}^{-1} A)^{-1} \quad (13)$$

3.5.2. Estimating Covariance Components (R and G)

If the design is perfectly balanced (that is, n_j all equal and the distribution of level 1 predictors

within each level 2 unit, there are closed-form formulae for estimating the variance-covariance parameters. When the design is unbalanced, iterative numerical procedures are used to obtain the estimates. Usually these procedures are based on maximum likelihood estimation techniques. Maximum likelihood (ML) estimates of G and R are found by maximizing the following log-likelihood function

$$l_{ML}(G, L) = -\frac{1}{2} \log|V| - \frac{N}{2} \log r^T V^{-1} r - \frac{N}{2} (1 + \log \frac{2\pi}{N}) \quad (14)$$

Where,

$$r = Y - A(A^T V^{-1} A)^T A^T V^{-1} Y$$

If the number of level 2 units, J is large then the estimates generated through maximum likelihood are approximately equal to estimates generated through restricted maximum likelihood (REML). REML estimates of the covariance components are based on residuals which are computed after estimating the fixed effects (11) by WLS or by GLS and are estimated based on maximizing a marginal likelihood. REML estimates take into account the degrees of freedom used in estimating the fixed effects when estimating the covariance components. REML estimates of G and R are found by maximizing the following log-likelihood

$$l_{REML}(G, R) = -\frac{1}{2} \log|V| - \frac{1}{2} \log|A^T V^{-1} A| - \frac{(N-p)}{2} \log r^T V^{-1} r - \frac{(N-p)}{2} [1 + \log \frac{2\pi}{(N-p)}] \quad (15)$$

Where;

$$r = Y - A(A^T V^{-1} A)^T A^T V^{-1} Y \text{ and } p = \text{rank}(A)$$

HLM generates REML estimates by default and uses the EM algorithm to maximize (11).

3.5.3. Estimating Random Effects (μ)

Random effects are estimated using shrinkage estimators. The estimates of random effects are generated according to the following:

$$\hat{\mu} = \hat{G} X^T \hat{V}^{-1} (Y - A\hat{\gamma}) \quad (16)$$

It is generally of interest to estimate the individual random coefficients. These can simply be obtained by substitution.

3.6. Asymptotic Properties of Estimators

We investigate consistency and asymptotic normality of each proposed estimator.

3.6.1 Consistency

Consistency means that when sample size is sufficiently large the estimator will be very likely to be very close to the actual parameter value. When an estimator converges in probability to the true value as the sample size increases, we say that the estimator is asymptotically consistent.

That is;

$$P \lim_{n \rightarrow \infty} (\hat{\beta}) = \beta \quad (17)$$

This means that as the sample size increases, the distribution of $\hat{\beta}$ degenerates: in the limit, the only possible realization of $\hat{\beta}$ is β .

(i) Weak consistency

Consistency is sometimes referred to as weak consistency. It allows the sample size to increase without bound. A sequence of estimators $\hat{\beta}$ is said to be weakly consistent estimator of β if

$$\hat{\beta} \xrightarrow{P} \beta$$

This means that the distribution of the estimates become more and more concentrated near the true value of the parameter being estimated, so that the probability of the estimator being arbitrarily close to β converges to one.

Definition: Estimator $\hat{\beta}$ is a (weakly) consistent estimator of the parameter β if, for every $\varepsilon > 0$ and every $\beta \in \Theta$, where ε is arbitrary positive small number. Then,

$$\lim_{n \rightarrow \infty} \Pr\{|\hat{\beta} - \beta| < \varepsilon\} = 1 \quad (18)$$

Informally, as the sample size becomes infinite (and the sample information becomes better and better), the estimator will be arbitrary close to the parameter with probability, an eminently desirable property.

An equivalent statement to this is for every $\varepsilon > 0$ and every $\beta \in \Theta$, a consistent estimators $\hat{\beta}$ will satisfy

$$\lim_{n \rightarrow \infty} \Pr\{|\hat{\beta} - \beta| < \epsilon\} = 0 \quad (19)$$

(ii) Strong Consistency

When an estimator converges almost surely to the true value, then the estimator is said to be strongly consistent. A sequence estimator $\hat{\beta}$ is said to be strongly consistent estimator of β if

$$\hat{\beta} \xrightarrow{a.s} \beta$$

Clearly, strong consistency implies weak consistency since almost surely (a.s.) convergence implies convergence in probability but not vice versa.

3.6.2. Asymptotic normality

The distribution of estimators are said to be asymptotically normal if, as the sample size increases, the distribution of the

estimators approaches a normal distribution. Mathematically, we say an estimate $\hat{\beta}$ is asymptotically normal if $\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N(0, \delta^2)$. Under asymptotic normality, the estimator not only converges to the unknown parameter, but it converges fast enough, at a rate $\frac{1}{\sqrt{n}}$.

4. Results and Interpretation

The data was analyzed using Stata software version 11 to determine the factors that influence crime rate in Kenya. HLM regression was used to estimate the parameters and then interpretation was made. Crime rate was disaggregated into crime against person and crime against property.

Table 1. HLM coefficients of fixed effects for crime against person.

Crime Rate	Coef.	Std. Err.	Z	P> Z	[95% Confidence Interval]	
					Lower	Upper
Intercept	10.11775	14.21819	0.71	0.004*	17.98489	37.7494
Poverty rate	0.95542	0.17758	5.38	0.000*	0.6073664	1.303476
Unemployment Rate	0.526087	0.1087708	4.84	0.000*	0.3129002	0.739274
School enrolment	-0.302063	0.1647508	1.83	0.067	-0.0208426	0.6249686
Young men(15-29 years)	0.239432	0.1473421	1.62	0.006*	0.05045783	0.5288657
Conviction Rate	-0.277573	0.1178335	1.86	0.018*	-0.0466233	0.508522
Clear-Up Rate	-0.870728	0.4126044	-2.11	0.035*	-1.679418	-0.062038
Probability of arrest	-2.653541	1.1685032	-1.27	0.023*	-4.943764	-0.363129
Population Density	0.0413675	0.1283256	1.22	0.001*	0.0162306	0.0664576
Police	-0.0296956	0.0084909	3.50	0.000*	0.0130538	0.0463374

4.1. Interpreting the Variables for Crime Against Person

From table 1 above for estimation of fixed effects, the parameter estimates for the variables Poverty Rate (0.95542), Unemployment rate (0.526), Young men (15-29) years (0.2394), Conviction Rate(-0.278), clear-up rate (-0.87) Probability of arrest (-2.65), population Density (0.0414) and police (-0.0297) are all significant at 95% confidence Interval in accounting for increase in Crime Rate against person with their expected signs. The variable school enrolment rate, with the expected sign in the coefficient (-0.302) was not significant in explaining the crime rate against person in the provinces.

Table 2. HLM estimates for variance components for crime against person.

	Estimate.	Std. Err.	[95% Confidence Interval]	
			Lower	Upper
Identity (Province)	19.71114	5.483453	11.42656	34.00228
Variance (residual)	7.328175	0.4217241	6.546522	8.203157

From table 2 above for estimation of the variance components, the first section of the table is labeled Identity (province), meaning that these are random variables at the second level (country) which gives the random intercept and is estimated as 14.82 with standard error 3.12. We have only

one random effect at the country level. The row labeled variance (Residual) displays the estimated standard deviation of the overall error term. This is the variance of the level-one errors, that is at the first level, the residuals and is estimated as 7.32 with standard error 0.4217. This implies there is a huge variation in crime rates against person between provinces.

Table 3. Summary of model fit for crime against person.

Log likelihood	Column2
Full Model	-615.7701
Model with Intercept Only	-617.3546
Chi-square	
LR (df=10)	165.3546
P-value	0.0000

According to table 3 above for the summary of the model fit, the overall model works well in explaining crime rate against person. The log-likelihood value of the model with the intercept only is -617.3546 while the full model is -615.7701. The reduction shows that the model is better at explaining increase in crime rate. An LR test comparing the model with one-level ordinary linear regression is provided and is highly significant for these data with the value 165.12, p value = 0.0000.

4.2. Interpreting the Variables for Crime Against Property

From table 4 above for estimation of fixed effects, the parameter estimates for the variables Poverty Rate (0.127), Unemployment rate (0.0414), School enrolment (-0.817), Probability of conviction (-0.126), Probability of arrest (-4.32), population Density (0.114) and police (-0.0054) are all significant at 95% confidence Interval in accounting for

increase in Crime Rate against property with their expected signs. The variables young men (15-29 years) with coefficient 0.227 and clear-up rate with coefficient -1.174 were not significant in explaining the crime rate against property in the provinces although they had the expected sign in the coefficient.

Table 4. HLM coefficients of fixed effects for crime against property.

Crime Rate	Coef.	Std. Err.	Z	P> Z	[95% Confidence Interval]	
					Lower	Upper
Intercept	56.71834	20.95356	2.71	0.000	15.65013	97.78656
Poverty rate	0.1269488	0.267851	0.47	0.003	-0.3980295	0.6519271
Unemployment Rate	0.4138228	0.166533	2.48	0.013	0.0874239	0.7402216
School enrolment	-0.817401	0.1252243	1.24	0.001	-0.3230138	1.3117891
Young men(15-29 years)	0.2274576	0.226324	1.00	0.316	0.216456	6.2695743
Conviction Rate	-0.125609	0.1799497	0.70	0.484	-0.2268339	0.4785558
Clear-Up Rate	-1.174315	0.6309484	-1.86	0.063	-2.410946	0.06232633
Probability of arrest	-4.322889	1.7880862	1.42	0.016	-4.183052	8.2747352
Population Density	0.1139867	0.1952754	0.28	0.000	0.2705476	4.9562314
Police	-0.005381	0.0129583	0.41	0.000	-0.0200596	0.0307358

Table 5. HLM estimates for variance components of crime against property.

	Estimate.	Std. Err.	[95% Confidence Interval]	
			Lower	Upper
Identity (Province)	25.94573	7.307859	14.93885	45.06245
Variance (Residual)	11.23459	0.6465346	10.03626	12.57601

The estimation table 5 shows the estimated variance components. The first section of the table is labeled Identity (province), meaning that these are random variables at the second level which gives the random intercept and is estimated as 25.95 with standard error 7.31. We have only one random effect at the country level. The row labeled variance (Residual) displays the estimated standard deviation of the overall error term. This is the variance of the level-one errors that is at the first level, the residuals and is estimated as 11.23 with standard error 0.647. This implies there is a huge variation in crime rates against property between provinces.

Table 6. Summary of model fit.

Log likelihood	Column2
Full Model	-682.25787
Model with Intercept Only	-749.50198
Chi-square	
LR (df=10)	146.58
P-value	0.0000

From table 6 for the summary of the model fit, the overall model works well in explaining crime rate against property. The log-likelihood value of the model with the intercept only is -749.50198 while the full model is -682.25787. The reduction shows that the model is better at explaining increase in crime rate. An LR test comparing the

model with one-level ordinary linear regression is provided and is highly significant for these data with the value 146.58, p value = 0.0000.

5. Conclusions

In this paper, we proposed a mixed effects regression model, that is a Hierarchical Linear Model to model crime rate and derive the estimators of the proposed model and further discuss the asymptotic properties of the proposed model. In order to test for the practicality of the proposed model, we estimate a crime equation using a panel dataset of provinces in Kenya for the period 1992 to 2012, employing the REML estimator. In this analysis, we use a provincial dataset; and explicitly consider factors that influence crime rates. Furthermore, instead of using the overall crime rate to measure the level of criminal activity, we separate the crime measure into property crimes and crimes against the person. The main conclusion of this paper is that all the crime rates considered differ significantly between provinces and hence crime should be studied at smaller level in order to understand it better.

6. Recommendations

6.1. Recommendations For Government and Policy Makers

In light of the results of this study, we have various recommendations to prevent the upsurge of crime in Kenya. The recommendations may help the government in formulation of policies that can be appropriate in crime prevention and move away from tradition way of crime reduction to preventive way. Increase in conviction rates in the long run appear to deter crime, hence the sure way to sustain it is by strengthening of the judicial system in Kenya which ensures that persons committing crimes are convicted. Police should make more arrests whenever crime is committed so as

to reduce crime levels in the locality. The community at large should also be educate the youth and occupy them with work.

6.2. Recommendations for Further Research

Statistical modeling of crime is a new field of study in the developing countries, this is due to the fact that most of the literature on the same in the said countries has not been exploited and Kenya is not an exception. Much of the Literature used is from the countries in the west and a few from East. Therefore, there is need for more research on the field particularly in Kenya so as to fight crime through informed framework. For further studies, we recommend the use Hierarchical Linear Models to analyze crime determinants at even lower levels like Districts, locations and villages in order to better reduce crime at all levels effectively.

Acknowledgements

I would like to thank my family for the financial support and encouragement and my supervisors Dr. Chelule and Dr. Mung'atu for their guidance in undertaking this study.

Abbreviations

HLM: Hierarchical Linear Model
REML: Restricted Maximum Likelihood
ML: Maximum Likelihood
LR: Likelihood Ratio
KNBS: Kenya National Bureau of Statistics

References

- [1] Asteriou, Dimitrios; Hall, Stephen G. (2011). "Misspecification: Wrong Regressors, Measurement Errors and Wrong Functional Forms". *Applied Econometrics* (Second ed.). London: Palgrave MacMillan. pp. 172–197.
- [2] Becker, G. (1968). Crime and Punishment: An Economic Approach. *The Journal of Political Economy*, 76/2, 169-217.
- [3] Brooks, Chris (2014). *Introductory Econometrics for Finance* (3rd Ed.). Cambridge: Cambridge University Press. p. 461. ISBN 9781107661455.
- [4] C.Cornwell and N.Trumbul. (1994). Estimating the Economic Model of Crime with panel Data. *The Review of Economics and Statistics*, Vol. 76.No.2(May,1994), 360-366.
- [5] Ehrlich. (1999). Crime,Punishment,and the Market for Offenses. *The Journal of Economic Perspectives*, Vol.10.1, 43-67.
- [6] Entorf and Spengler. (2002). *Crime in Europe: Causes and Consequences*. Springer.
- [7] Gelman, A. and J. Hill. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York, NY: Cambridge University Press, 2007.
- [8] Gujarati, Damodar N.; Porter, Dawn C. (2009). "Panel Data Regression Models". *Basic Econometrics* (Fifth international ed.). Boston: McGraw-Hill.
- [9] Hosmer, David (2013). *Applied logistic regression*. Hoboken, New Jersey: Wiley. ISBN 978-0470582473.
- [10] Howell, David C. (2010). *Statistical Methods for Psychology, 7th ed.* Belmont, CA; Thomson Wadsworth. ISBN 978-0-495-59786-5.
- [11] Levitt, S. (1996). The Effect of Prison Population Size on Crime Rates: Evidence from Prison Overcrowding Litigation. *The Quarterly Journal of Economics*, 111/2, 319-351.
- [12] Marselli and Vannini. (1997). Estimating a Crime Equation in the Presence of Organized Crime:Evidence from Italy. *International Review of Law and Economics* 17, 89-113.
- [13] Schafer. (1976). Fear of Crime and Criminal Victimizations:Gender-based Contrasts. *Journal of Criminal Justice*, Vol.34, 285-301.
- [14] Zuur, A. F., Ieno, E. N., Walker, N., Saveliev, A. a., & Smith, G. M. (2009). *Mixed effects models and extensions in ecology with R*. Statistics. New York, NY: Springer. doi:10.1007/978-0-387-8745.
- [15] Levit, S.(2001). Alternative startegies for Identifying the Link between Unemployment and Crime. *Journal of Quantitative Criminology*, 17/4,377-390.
- [16] Orsagh. (1979). Empirical Criminology: Interpreting Results Derived From Aggregate Data. *Journal Of Research in Crime and Delinquency*, 294-306.
- [17] Witte, A. D. (1980). Estimating the Economics Model of Crime with Individual Data. *The quarterly journal of economics*, Vol. 94, No. 1., 57-84.
- [18] Ehrlich, I. (1973). Participation in Illegitimate Activities: A Theoretical and Empirical Investigation. *The Journal of Political Economy*, 81/3, 521-565.
- [19] Becker, G. (1993). Noble Lecture: The economic way of looking at behaviour. *Journal of political econmy* 101, 385-409.
- [20] Baltagi and Akom. (1990). On Efficient Estimation with Panel Data: An Empirical Comparison of Instrumental Variables Estimators. *Journal of Applied Econometrics*, Vol.5,No.4, 401-406.John Wiley & Sons.
- [21] Andreoni. (1991). Reasonable Doubt and Optimal magnitude of fines: should the punishment fit crime? *RAND Journal of Economics*, 22 (3), 385-395.
- [22] Bryk, A.S. and S.W. Raudenbush (1992) *Hierarchical Linear Models, Applications and Data Analysis Methods*. Newbury Park, CA: Sage.
- [23] Raudenbush, s. W., Bryk, A. s., Cheong, Y. F., Congdon, R., & Du Toit, M. (2011). *HLM 7: Hierarchical linear and nonlinear modeling*. Lincolnwood, IL: scientific software International.