

Estimating Victimization Rates, Trends and Risk Factors in Palestine

تقدير معدلات واتجاهات وعوامل المخاطرة للإيذاء في فلسطين

Mahmoud Okasha* & Nisreen Nouraldin

محمود عكاشة، ونسرين نور الدين

Department of Applied Statistics, Al-Azhar University, Gaza, Palestine

*Corresponding Author: E-mail: m.okasha@palnet.com

Received: (23/12/2012), Accepted: (26/1/2014)

Abstract

This paper examines different risk factors that are most influential in identifying potential victims who have the greatest risk to be victimized in Palestine. The goal of this is to introduce a powerful statistical model that can be used in estimating the prevalence rates, trends and risk factors for a specific phenomenon such as victimization in Palestine as well as in other similar phenomena. The risk factors for being victimized in Palestine including many demographic, economic and social variables have been studied. The goal is to find those significant risk factors and identify the characteristics of the persons who are prone to tangible losses. The data used in the analysis cover all the Palestinian Territories including the West Bank, Gaza Strip and East Jerusalem as of 2008 and has been gathered by the Palestinian Central Bureau Of Statistics. Logistic regression model has been used in the analysis, and revealed that the significant influential factors on victimization are sex and owning a car. The effects of each factor, its odds ratio and risk ratio are examined. The model has been applied to estimate the probability of a person to be victimized and to classify the persons who are vulnerable to be victimized, and succeeded in correctly classifying 64% of people who have really fallen victims and 96% of people who are unbeatable to crime. The general percentage of correct classification was 93.5%.

Key words: Classification, Logistic Regression Model, Odds Ratio, Victimization

ملخص

تبحث هذه الورقة في عوامل المخاطرة المختلفة للإيذاء في فلسطين للتعرف على العوامل الأكثر تأثيراً في تحديد ضحايا الإيذاء المحتملين والمفيدة في التعرف على ضحايا الجريمة في فلسطين، كما تقدم الورقة أسلوباً إحصائياً متقدماً كنموذج لتحليل بيانات الاستبانات يمكن من خلاله تقدير معدلات واتجاهات ظاهرة معينة والتعرف على عوامل المخاطرة لها مثل ظاهرة الإيذاء في فلسطين، وقد تم دراسة معنوية العديد من عوامل المخاطرة على الإيذاء في فلسطين مثل عوامل المنطقة الجغرافية والمتغيرات الديموغرافية وبعض المتغيرات الاقتصادية والاجتماعية. وتهدف الورقة من خلال ذلك إلى التعرف على عوامل المخاطرة المعنوية والمؤثرة على الإيذاء وتحديد خصائص الأشخاص الأكثر عرضة لخسائر ملموسة. وقد بنيت الدراسة على تحليل إحصائي لبيانات عن المجتمع الفلسطيني تغطي كل الأراضي الفلسطينية بما فيها الضفة الغربية وقطاع غزة والقدس الشرقية خلال العام ٢٠٠٨، حيث تم الحصول على هذه البيانات من الجهاز المركزي للإحصاء الفلسطيني، وتم استخدام نموذج الانحدار اللوجستي الإحصائي في التحليل والكشف عن أكثر العوامل المعنوية إحصائياً تأثيراً على الإيذاء في فلسطين وتبين من التحليل أنها تمثلت في متغيري الجنس وامتلاك الأسرة لسيارة خاصة، كما تم فحص آثار كل عامل من عوامل المخاطرة، ونسب الخلاف والمخاطرة لكل متغير، إذ يمكن استخدام النموذج المقترح في تقدير احتمال أن يكون شخص ما ضحية للإيذاء خلال عام معين، وكذلك التعرف على فئة الأفراد الأكثر عرضة ليكونوا ضحايا محتملين، ونجح النموذج في تصنيف الأفراد الذين وقعوا ضحايا للإيذاء حقاً في عينة الدراسة بشكل صحيح بنسبة ٦٤٪ من مفردات عينة الدراسة، ونسبة ٩٦٪ من الأشخاص الذين لم يتعرضوا للإيذاء في فلسطين، وبشكل عام كانت نسبة التصنيف الصحيحة هي ٩٣,٥٪.

Introduction

Crime and victimization became one of the major problems in Palestine in recent years. Victimization rates are severely increasing. The Israeli occupation to the Palestinian land and the violence of the Israeli settlers and their attacks on Palestinian people and homes are the main sources of victimization in Palestine. Moreover, the increasing high unemployment and poverty rates among the Palestinian population seem to be the main reasons for the increasing trend of victimization in Palestine. However, victimization as a new prevailing phenomenon did not receive attention by Palestinian researchers. Yet, victimization rates and risk factors of victimization have never been studied in Palestine. The Palestinian Central Bureau of Statistics (PCBS) published many data sets on this phenomenon but these data sets have never been analyzed to

shed the lights on this problem. The importance of this study derives from the scarcity of quantitative researches that handled the victimization phenomenon using sophisticated statistical methodology both worldwide and in Palestine. The aim of this paper is to estimate the victimization rates in Palestine, by applying various appropriate statistical models in order to make preventive and treatment plans to limit the spread and increase in victims and control crime causes. Moreover, this study applies one of the advanced statistical modeling techniques on one of the most important issues in the Palestinian society that can be applied in similar data sets on different topics that exist in the PSBS.

The concept of *victim* dates back to ancient cultures and civilizations. Over the centuries, the word *victim* came to have additional meanings. During the founding of victimology in the 1940s, victimologists such as Mendelson (1963) and von Hentig (1948) tended to use textbook or dictionary definitions of victims as hapless dupes who instigated their own victimizations. This notion of "victim precipitation" was vigorously attacked by feminists in the 1980s, and was replaced by the notion of victims as anyone caught up in an asymmetric relationship or situation.

Research Problem

The victim is one of crime action pillars that consist of criminal act, crime and the victim. Victimization has a role in specification and shaping the criminal act. The research problem of this paper is how to estimate victimization rates, trends and risk factors using the available data from the PCBS in order to explore existing specific characteristics of the victims in the Palestine society. The importance of this study comes from the little and scarcity of quantitative scientific research on victimization in the literature that applies advanced statistical methods both worldwide and in Palestinian society. The goal is to make scientific investigation and practical action to find the appropriate recommendations towards taking preventive measures and treatment plans to limit the spread of victimization and control crime causes in Palestine.

The Data

The data of this study is a survey covering all the Palestinian Territories including the West Bank, Gaza Strip and East Jerusalem and has been conducted by the Palestinian Central Bureau Of Statistics (PCBS, 2008) based on a household sample survey during the period from 04/10/2008 until 31/12/2008. It provides basic indicators on various aspects of victimization, including households victimized by criminal offenses, type of criminal offense, tangible losses of crimes, Crime location, Perpetrator, Crime reporting, Reported crime underwent legal proceedings, Which party prone to tangible losses. A special questionnaire was designed and recommendations in the field of victimization statistics while taking the Palestinian particularities into account. The survey covered 10260 sampling units of Palestinian households from all Palestinian governments. The survey's methodology was designed taking into account the Palestinian conditions, international standards, data processing requirements and the comparability of outputs with other related surveys conducted in the Palestine.

The survey data used in this study has been provided by the PCBS under a special agreement of cooperation between the PCBS and Al-Azhar University- Gaza for the purposes of teaching and scientific research. One limitation on the use of the data is that the survey does not cover victimization in both West Bank and Gaza Strip which came as a result of the Palestinian separation and the victimization due to the peoples political views since the survey has been conducted in 2008 and after the Palestinian separation in 2007. This is because this type of victimization is not considered a type of criminal offence in both parts of Palestine, and in many cases it is not allowed to report such cases. Thus, in practical use of the results, it should be noted that the overall victimization rates are underestimated since the data are constrained by criminal acts only and does not cover cases of political victimization.

Historical Background

A few studies that employed statistical models on victimization data had been found in the literature all over the world (Karmen, (2003). The

majority of those only used simple data summary and presentations and in the best cases some indicators like averages, rates and proportions had been estimated using simple formula (Engelhardt, *et al.*, 2008, and Hickey, 2006). We now summarize some of those papers particularly those employed statistical methods in the analysis. In a study of Stein (2009), six groups of countries that highlights the importance of distinguishing between areas of the world beyond only industrial and non-industrial categorizations, during the four years: 1989, 1992, 1996, and 2000. The sample size for each country ranges from approximately 1,000 to 2,000 respondents per survey. Within this research, the dependent variables are dichotomous, scored 1 if the respondent has been victimized and scored 0 if not. Hierarchical Linear Models took into consideration the layered or nested nature of the data, nesting respondents within the country in which they reside. Marcotte and Markowit (2009) in their paper explored the relationship between trends in treatment for mental illness and violent crime. The researchers tried to characterize the behavioral mechanisms for these relationships by summarizing important syndromes and how they might contribute to behaviors leading to criminal acts and also increase risk of victimization.

In a paper by Tyler *et al.*(2004), the risk factors associated with the likelihood of being sexually victimized by a stranger friend/acquaintance since being on the street were examined among 372 homeless and runaway youths. Young people were interviewed on the streets and in shelters by outreach workers using a systematic sampling strategy. Youths who had been engaged in higher risk behaviors were expected to be at greater risk for sexual victimization by both known and unknown assailants. Results indicated that for females running from home for the first time at an earlier age were associated with sexual victimization by both a stranger and friend/acquaintance. However, engaging in deviant subsistence strategies, survival sex, and grooming predicted being sexually victimized by a friend/acquaintance. For males, survival sex and grooming predicted stranger sexual victimization, whereas sexual orientation was associated with sexual victimization by a

friend/acquaintance. Overall, 35% of the sample had been sexually victimized.

Koo and Pierre (2003) conducted a study to estimate the prevalence of violent victimization in a 30 days period among a sample of 900 street recruited heroin users in Miami- Dade County, Florida, to identify the risk factors for violent victimization among this drug group, and examine two different types of violent victimization (robbed and injured) and analyze whether risk factors vary among these different types of victimization. Roodman (2000) conducted another study to examine two competing models of sexual victimization that examined the path between child abuse and later sexual victimization. Structural equation modeling was used to examine two competing models of sexual victimization. The study was based on a sample of 276 college students taking introductory psychology.

Gaviria and Pagés (1999) in their paper used the Latino barometer to study the patterns of crime victimization in Latin America. The Latino barometer is a public opinion survey covering 17 Latin American countries. The survey has been regularly conducted since 1996. Roughly, 1,500 individuals have been interviewed in each country each year. The sampling method varies slightly from country to country. Straus (1986) focused on criminal homicide, defined as death due to injuries illegally inflicted by another person with intent to injure or kill by any means. Determination that a death was criminal homicide was based on the results of investigation by the Los Angeles Police Department. Demographic characteristics of victims and perpetrators and situational characteristics of the homicide were obtained from confidential police files. Sampson (1985) used the National Crime Survey victimization data from 1973-1978 to examine the effects of neighborhood characteristics and extent of urbanization on rates of theft and violent personal victimization. The results underscore the importance of urbanization and the physical environment in predicting victimization risk.

Methodology

Since our main interest is on the response variable “households victimized by criminal offense” is binary and we have many independent binary and categorical variables of interest in the survey, the main statistical model which can be used in the analysis of the victimization dataset in this paper is the logistic regression model. Other models that can be used for the analysis of a categorical response variable such as discriminant analysis require numeric independent variables and have a multivariate normal distribution. Those conditions do not exist in our data set. Thus, logistic regression is the ideal model for our analysis since the main goals of our study are to determine the best subset of independent variables, among various binary and categorical independent variables, that best predict group membership of cases using data of known groups of a dichotomous dependent outcome variable.

Concepts and Definitions

In this section we provide some basic definitions of the legal and social concepts that are used in the present paper and as had been defined by PCBS (2008).

Assault

The term assault refers to physical attack against persons, but excludes indecent assault. Some criminal or penal codes distinguish between aggravated and simple assault depending on the degree of resulting injuries.

Crime

Crime is defined as any act involving violation of laws or public rights duties towards the state or society in general.

Crime Location

This term refers to the place where the crime took place.

Household

It refers to a group of at least one person living together who make common provisions for food or other essentials for living. Households members may be related, unrelated or a combination of both.

Number of Households (n)

Since the household is the sampling unit, the number of households is the sample size. Sometimes, it is weighted sample size.

Perpetrator

The person violating effective laws by undertaking criminal events against other persons or their properties is referred to as perpetrator.

Physical Harm

All losses that a person may suffer during the crime that took place in the last 12 months prior to the survey, which resulted in wounds, murder, malformation or disability is referred to as physical harm.

Properties

All movable and fixed assets belonging to the individuals (household members) regardless of whether they are inside or outside the house are termed in this paper as properties.

Robbery

Illegally breaking into the property of somebody with the intention to commit a crime is referred to as robbery.

Theft

The term theft here refers to the removal of property without the property owner's consent. Theft excludes burglary and house breaking; such as the theft of a motor vehicle, shoplifting and other minor offenses like pilfering and petty theft may or may not be included as thefts.

Victim

The person affected by an offense or loss or prey to catastrophic, criminal or brutal events is termed here as victim. Any person who was offended and whose properties were partially or totally affected by a criminal act or incident is classified as a victim.

The Logistic Regression Model

The logistic regression model (LRM) is the natural complement of ordinary linear regression whenever the response is a binary variable (Agresti, 2007). For a binary response variable Y with two measurement levels (dichotomous) and explanatory variables x_1, x_2, \dots, x_k with k denotes the number of predictors, the model has a linear form for the *logit* of probability (log odds) as follows:

The form of the logistic model formula is:

$$\begin{aligned} \text{logit}[\pi(\underline{x})] &= \text{logit} \left[P(Y = 1 | X_1, \dots, X_k) \right] = \log \left[\frac{\pi(\underline{x})}{1 - \pi(\underline{x})} \right] \\ &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k \end{aligned}$$

where the *odds* ($Y=1$) = $\frac{\pi(\underline{x})}{1 - \pi(\underline{x})} = \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)$.

The logistic regression model implies the following formula for the probability $\pi(\underline{x})$:

$$\pi(\underline{x}) = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)}$$

The above formula shows that logistic regression is really just a standard linear regression model, once we transform the dichotomous outcome by the logit transform. This transform changes the range of $\pi(\underline{x})$ from 0 to 1 to $-\infty$ to $+\infty$, as usual for linear regression. Thus, the

logits can take on any values between $-\infty$ to ∞ while $\Pr(Y=1)$ can only take on values between 0 and 1.

An adjusted odds ratio is an odds ratio comparing two categories of the variable after controlling for the other variables in the model (Moorman and Carr, 2008). For example, an adjusted odds ratio comparing two categories of the variable X_1 is

$$OR_{X_1=1 \text{ vs } X_1=0} = \frac{\text{Odds}(Y = 1 / X_1 = 1, X_2, \dots, X_k)}{\text{Odds}(Y = 1 / X_1 = 0, X_2, \dots, X_k)} = \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 + \hat{\beta}_2 X_2 + \dots + \hat{\beta}_k X_k}}{e^{\hat{\beta}_0 + \hat{\beta}_2 X_2 + \dots + \hat{\beta}_k X_k}} = e^{\hat{\beta}_1}$$

and its $(1 - \alpha) * 100\%$ confidence interval is

$$e^{\hat{\beta}_1 \pm Z_{1-\alpha/2} S_{\hat{\beta}_1}}$$

The logistic procedure fits linear logistic regression models for dichotomous variables by the method of maximum likelihood estimation. Let $-2 \log L_A = \log$ -likelihood statistic of model A with “p” predictors and $-2 \log L_B = \log$ -likelihood statistic of model B with “k” predictors and $k > p$.

Then the likelihood ratio Chi-square, G^2 , is

$$G^2 = (-2 \ln L_A) - (-2 \ln L_B) \sim \chi^2_{k-p}$$

If Model A is the model with intercept only, then G^2 plays the role of the overall F, testing $H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$ with “k” degrees of freedom.

If Model A has “p” predictors and model B has “k” predictors, then G^2 plays the role of the (multiple) partial F with “k – p” degrees of freedom.

The Wald Chi-Square Test is

$$W = \frac{\hat{\beta}_i^2}{\text{SE}(\hat{\beta}_i)^2}, \quad i = 1, \dots, k$$

It tests $H_0 : \beta_i = 0$ for $i = 1, \dots, k$ vs $H_A : \text{Not } H_0$.

The Salient Features of the Data

Victims of Criminal Offenses at the Households Level

Simple descriptive analysis of the data showed that 7.5% of the Palestinian households were exposed to criminal offenses during the year 2008, 5.8% of the households were victimized in the West Bank households and 10.9% of them in Gaza Strip. The analysis showed that the percentage of households in Palestine that were exposed to theft (excluding vehicles theft) is 2.0%, 6.1% for vehicle theft or part of it, 1.7% property damage, 0.8% threat, and 0.9% assault. In general the data revealed that the percentage of households victimized by criminal offenses in Gaza Strip is higher than that in the West Bank, except for households victimized by vehicle theft or part of it: 6.2% in the West Bank and 5.9% in the Gaza Strip, robbery or theft attempt: in West Bank 0.7% and in the Gaza Strip 0.3%. The results showed that 2.4% of the households in Palestine were exposed to Israeli soldiers or settlers' harassment and assault, compared with 7.1% in 2004. Further statistical indicators can be found in NouralDin (2012).

Victims of Criminal Offenses at the Individual Level

Type of Criminal Offense

Analysis of the survey data revealed that 33.9% of the individual victims of criminal offenses in the Palestine were exposed to theft and attempted theft (these percentages are 42.1% in the West Bank and 26.3% in the Gaza Strip). Moreover, 18.4% of the individual victims of criminal offenses were exposed to threat\ assault (12.3% in the West Bank and 24.0% in the Gaza Strip). The percentage of individual victims of criminal offenses in the Palestine exposed to property damage is

18.3%. (13.3% in the West Bank and 23.0% in the Gaza Strip).The survey results also revealed that 27.8% of the victimized Palestinian persons were a result of the Israeli Soldiers or Settlers Harassment and Assault (30.1% in the West Bank and 25.6% in the Gaza Strip).

Crime Location

The results indicated that 45.7% of criminal offenses in the Palestine took place inside the house, 26.5% nearby their house, 18.7% in another place inside locality, and 9.1% outside the locality (See figure 1). Criminal offenses occurring inside the house were higher in the Gaza Strip (52.2%) than in the West Bank (38.8%), and criminal offenses that took place outside the locality were lower in the Gaza Strip (3.7%) compared with the West Bank (14.8%).

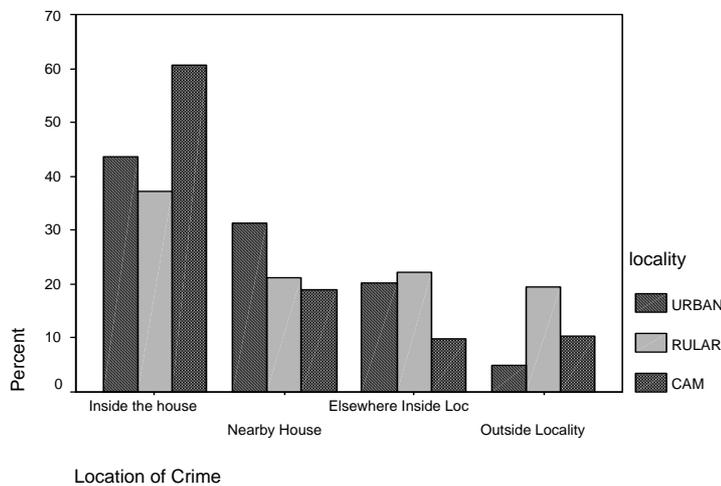


Figure (1): Percentage Distribution of Victimized Persons by Location of Last Crime and Type of Locality during Last 12 Months, in Palestine as of 2008.

Perpetrator

The results showed that 33.2% of criminal offenses against persons in the Palestine were committed by Israeli soldiers or settlers (41.7% in

the West Bank and 25.1% in the Gaza Strip). Figure 2 shows also that about 3.8% of these criminal offenses were committed by one of the relatives (6.5% in the West Bank and 1.4% in the Gaza Strip).

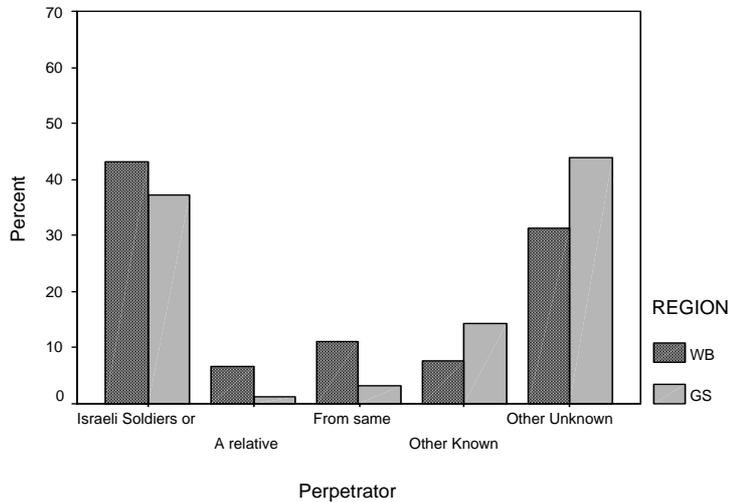


Figure (2): Percentage Distribution of Victimized Persons by Perpetrator of Last Criminal Offense and Region in Palestine during Last 12 Months, as of 2008.

Physical Harm and Tangible Losses of Criminal Offenses

The percentage of criminal offenses that caused physical harm in Palestine is 12.2% (13.1% in the West Bank compared with 11.5% the Gaza Strip). The results show that 30.4% of criminal offenses against persons in Palestine caused tangible losses of more than 1000 Jordanian Dinars (distributed as 34.2% in the West Bank and 26.9% in the Gaza Strip). In about 72.7% of criminal offenses cases against persons in Palestine, as indicated in figure 3 below, the victims were subjected to tangible losses, compared with 85.8%, 88.7% and 78.6% in the years 1996, 1999 and 2008 respectively.

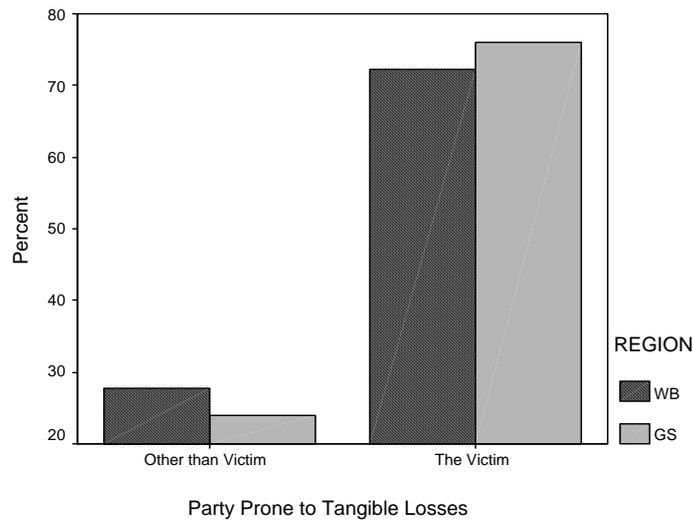


Figure (3): Percentage Distribution of Victimized Persons by Party Prone to Tangible Losses of the Last Criminal Offense in Palestine Last 12 Months, as of 2008.

Fitting the Logistic Regression Model

Logistic regression analysis has been applied to analyze the data of this study. This model is suitable for data where the researchers aim at identifying the variables that can predict group membership for known groups, particularly when the dependent variable is dichotomous. Further, logistic regression reveals the percent of the variance in the dependent variable accounted for by the independent variables. In addition, logistic regression can establish a hierarchy of significance for individual independent variables in the overall model, as well as explaining interaction effects. There are several advantages for using logistic regression in the statistical analysis of the current study. First, the rigid assumptions of other forms of regression do not apply to logistic regression. For example, there is no assumption of a linear relationship between the dependent variable and the independent variables. Also, there is no assumption that the dependent variable is normally distributed

in the population. In addition, there is no assumption of homogeneity of variance. Accordingly, it is not required that the dependent variable be homoscedastic for each level of the independent variables. Moreover, logistic regression does not assume that the error terms are normally distributed. Lastly, there is no requirement that independent variables be interval or unbounded.

The non-parametric version of logistic regression analysis was used in this study since all independent variables were categorical (such as region, sex, job, owning a car, which party prone to tangible losses, place of the crime, attempting break the house and reception a threat call) and parametric tests require interval data. The dependent variable, which measured the outcome of victim versus not victim, was a discrete variable. As such, ordinary least squares regression could have been used to fit a linear probability model. However, because the linear probability model is heteroscedastic and could predict probabilities less than 1 or greater than 0, logistic regression was more appropriate to identify the factors that predict victims. It can also be used to estimate the probability of a person to be victimized.

The Models

In this paper, all the available independent variables were used in building the logistic regression model for the data. To examine the overall fit of the model, the model chi-square was computed. To examine the proportion of the variance in the dependent variable explained by the variance in the independent variables, the Cox and Snell R-squared and the Nagelkerke R-squared statistic were also computed. The logistic regression analysis revealed the factors that should be included in the model to classify the Palestinian households being victimized versus not victimized. There are many hypotheses which can be tested here. The most important hypothesis states that there is a significant relation between the variable Households victimized by criminal offense and some other independent variables. These hypotheses were tested using the Wald statistic, which is the square of the asymptotic t-statistic from the logistic regression analysis. In addition, a chi-square test for independence was used to test this hypothesis.

All hypothesized predictor variables were entered into a stepwise binary logistic regression model. In the initial analysis, a nine-predictor logistic model was fitted to the data to test the research hypothesis. The following nine predictor variables were used: Reasons for not reporting (cr11); Region; household job (HHocup); Is the household or any member of the household owned a car (Vs2); sex; which party prone to tangible losses (cr18); where did the crime happen (cr04); is there anything to indicate that somebody attempted to break the house (vs4); and did household or any member exposed to threat exception threat calls (vs7). The parameters of the model have been estimated and the final results are exhibited in table (1) below.

Table (1): Summary of Logistic Regression Analysis Results.

| Final Model | | | | | | | |
|----------------------------------------|--------|--------|------------|------------------|-------------------|----|--------|
| Predictor | b | SEb | Odds ratio | 95% CI | Wald's Chi-square | df | p |
| X ₁ | -4.046 | 0.9569 | 0.175* | 0.0027 to 0.1141 | 17.8782 | 1 | 0.0001 |
| X ₂ | 2.7343 | 0.8834 | 15.399 | 2.726 to 86.9775 | 9.5813 | 1 | 0.002 |
| Test | | | | | χ^2 | df | p |
| Hosmer & Lemeshow Goodness-of-fit test | | | | | 3.2246 | 5 | 0.665 |

*The inverse of this odd ratio 5.714 is more meaningful.

It should be noted that the intercept of the model b_0 was not significant. Therefore it has been omitted from the model and analysis has been conducted again without an intercept. Table (1) above displays the final results of stepwise logistic regression analysis with no constant term included. From this table we can see that the best fitted logistic model is as follows:

$$\text{Logit}(Y) = -4.046 X_1 + 2.7343 X_2$$

Where Y : Households victimized by criminal offense.

X_1 : family has its own car (0=no, 1=yes).

X_2 : Sex (1=male, 0=female).

According to the above model and the results of table (1), the log of the odds of Households victimized by criminal offense being a victim was positively related to sex with ($p < 0.05$) and negatively related to family has its own car ($p < .005$).

Variance Explained by the Model

To examine the proportion of variations in the dependent variable explained by the variance in the independent variables, the Cox and Snell R-squared and Nagelkerke R-squared statistics were used. The Cox and Snell $R^2 = 0.211$ and the Nagelkerke $R^2 = 0.478$, indicating that the model explained only 21.1% (using Cox and Snell R^2 coefficient) and 47.8% (using Nagelkerke R^2 coefficient) of the variations in the dependent variable due to the variations in the predictor variables. This indicates that there is still quite a large proportion of the variation in the dependent variables not explained by the predictor variables involved in the analysis. This result calls for inclusion of other predictor variables in the analysis some of them are not available in the questionnaire of the PCBS (2008) such as the educational level, economic status and material status of the victims and may be also the perpetrators themselves.

Goodness of Fit Tests

The Hosmer-Lemeshow (H-L) test is an inferential goodness of fit statistic used to assess the fit of a logistic model against actual outcomes, households victimized by criminal offense in our case (Hosmer and Lemeshow, 2000). The H-L test is used to test the null hypothesis that the fitted model is significant for the underlying data against the alternative hypothesis that the model is not significant. The test yielded a small (χ^2) test value and a large p-value [$\chi^2(5, N = 293) = 3.2246, p > 0.05$]. This

means that there is no evidence to reject the null hypothesis and indicates that the final model was a good fit of the data.

Statistical Tests of Individual Predictors in Overall Model

The Wald Chi-square statistic was used to test the statistical significance of the individual regression coefficients. According to Table (1), the variables X_1 and X_2 were significant predictors of the outcome of an incident of Households to be victimized by criminal offense ($p < 0.05$). Based on the chi-square statistics above of the two predictors in the final model, the Bs (coefficients) of those predictors are significantly different from 0. Thus the null hypothesis (model) is rejected. Table (1) shows regression coefficients, Wald statistics, and the coefficient of each of the significant predictors.

Odds Ratios

The odds ratios, as shown in Table (1), for the predictor of X_1 (OR = 0.175) revealed that when a family has its own car is almost six (1/0.175) times less chance than Households to be victimized by criminal offense than the family who doesn't have its own car. This means that 1 unit increase in family that has its own car increases the odds of being a Households victimized by criminal offense by a multiple of 0.175. Moreover, for the predictor of X_2 (OR = 15.399) revealed that the odds of male victims are almost fifteen times more than the odds of female victims. This means that a 1 unit increase in male victims increases the odds of being a Households victimized by criminal offense by a multiple of 15.399.

Validations of Predicted Probabilities

The overall model's percentage accuracy in classification is 93.52%, which is an improvement over the null model's percentage accuracy in classification of 91.47%. As may be seen in Table (2), the correct classification of cases that were victims was less accurate than the correct classification of cases that were not victims. The sensitivity level for accurately classifying victims was 64% with 16 out of 25 cases accurately classified. The specificity level was 96.27%, with 258 out of

268 cases accurately classified. The false positive rate, or the proportion of observations erroneously classified as victims (n = 10) over all cases classified victims (n = 26), was 38.5%. Therefore, the positive classification value of the model was 61.5%, which is the proportion of observations correctly classified victims (n = 16) over all observations classified victims (n = 26). The false negative rate - the proportion of observations erroneously classified not victims (n = 9) over all cases classified not victims (n = 267) was 3.37%. The negative classification value of the model was 96.6%, which is the proportion of observations correctly classified not victim (n = 258) over all observations classified not victim (n=267).

Table (2): Accuracy of Predictions of Households Victimized by Criminal Offenses in the Palestinian Society Using the Final Model.

| Observed | | Predicted | | |
|----------------------------|-----------|-------------------------------------------|-----------|--------------------|
| | | Households victimized by criminal offense | | Percentage Correct |
| | | victim | no victim | |
| Null model | victim | 0 | 25 | 0% |
| | no victim | 0 | 267 | 100% |
| Overall Percentage Correct | | | | 91,47% |
| Final model | victim | 16 | 9 | 64% |
| | no victim | 10 | 258 | 96,27% |
| Overall Percentage Correct | | | | 93.52% |

Examination for Multicollinearity

Strong correlations between independent variables can result in multicollinearity in logistic regression models. This can inflate the variances of the parameter estimates. When there are small or moderate sample sizes, multicollinearity can result in lack of statistical significance of individual independent variables even when the overall model has achieved significance. To test for multicollinearity, the diagnostic statistics of Tolerance and Variance Inflation Factor (VIF) in linear

regression were used. Tolerance and VIF values indicated that no multicollinearity exist among the independent variables included in the analysis of our data set.

Conclusion and Recommendations

Conclusion

The current paper examined eight factors as candidates for being most useful in classifying victims of crimes and their risk on the households to be victimized in the Palestinian society. Those factors include region, sex, job, owning a car, party prone to tangible losses, place of the crime, attempting to break the house and reception of a threat call. The current study found statistical evidence to support only two of these factors: sex and the status of owning a car. In other words men who do not own cars are the most exposed group to be victims of crime in the Palestinian society. This gives an indication that poor men are the most prone group to be victimized.

The model has been applied to classify the occurrence of persons to be victimized and succeeded in correctly classifying (64%) of people who have really fallen victims and (96%) of people who are vulnerable to crime. The general percentage of correct classification was (93.5%).

The paper introduces a new statistical technique that can be employed to analyze questionnaire data sets which involve an underlying qualitative phenomenon and many predictors that can be qualitative or quantitative. The PCBS collects and publishes many data sets regularly and such a model can be useful in the analysis of many of these data sets and their results can be of great benefits to the Palestinian society.

Recommendations

- Logistic regression proved to be a good statistical technique that can be very useful in the analysis of questionnaire data sets such as those regularly published by the PCBS.
- The victimization survey should be improved by the PCBS to include further questions related to educational level, economic status and material status of the victims and may be other questions for deeper analysis.
- A database including crimes and offenses that include full records on perpetrators should be established in Palestine.
- The data should be updated regularly to facilitate tracking of crime indicators in Palestine.
- Plans for improvement of the security services and handling of crimes particularly among the poor people and within poor areas in Palestine should be developed.
- Further studies that take account of other new factors to be included in the questionnaire and using other statistical tests and techniques should be conducted.
- Cooperation between PCBS and the Palestinian Ministry of interiority should be established to collect frequent data for crime prevention, particularly from criminals and victims records.
- The results of this papers indicated that males who belong to poor families are the most prone group in the Palestinian society to victimization; therefore, this group should be given more care and security measures and public awareness campaigns.

Acknowledgement

The data used in this study has been provided by the Palestinian Central Bureau of Statistics (PCBS) under a special agreement of cooperation between PCBS and Al-Azhar University- Gaza for the purposes of teaching and scientific research. The authors are thus very

grateful to the PCBS’ president and staff for their efforts and anticipation since without their kind cooperation this work would have been impossible to be conducted. The authors hope for continuing cooperation between PCBS and the Palestinian universities for the benefits of students, scientific research and the Palestinian society.

References

- Agresti, A. (2007). An Introduction to Categorical Data Analysis. John Wiley & Sons. Inc.
- Engelhardt, B. Rocheteau, G. & Rupert, P. (2008). “The Labor Market and Female Crime”. Chapter 4(1). 139 – 163
- Gaviria, Alejandro. & Pages, Carmen. (1999). “Patterns of Crime Victimization in Latin America (October 1999)”. IDB Working Paper No. 339. Available at <http://ssrn.com/abstract=1817205> or <http://dx.doi.org/10.2139/ssrn.1817205>.
- Hickey, Eric. (2006). Victims. In Serial murderers and their victims. Belmont. CA: Wadsworth Group.
- Hosmer, David & Lemshow, Stanley. (2000). Applied Logistic Regression. New York : Wiley.
- Karmen, Andrew. (2003). Crime Victims: An Introduction to Victimology. Wadsworth Publishing.
- Koo, Dixie. & Pierre, Karen. (2003). “The Prevalence & Risk Factors for Violent Victimization Among Street Recruited Heroin Users”. Paper presented at the annual meeting of the American Sociological Association. Atlanta. GA. Aug 16. 2003 Available at: http://www.allacademic.com/meta/p107620_index.html

- Marcotte, Dave. & Markowit, Sara. (2009). A cure for crime? Psycho-pharmaceuticals and crime trends."NBER Working Papers 15354. National Bureau of Economic Research. Inc.
- Mendelson B. (1963). "The Origin of the Doctrine of Victimology". *Excerpta Criminologica* 3:30
- Moorman, S.M. & Carr, D. (2008). Spouses Effectiveness as End-of-Life Health Care Surrogates: Accuracy. Uncertainty. and Errors of Overtreatment or Undertreatment. *The Gerontologist* 48:811. The Gerontological Society of America.
- NouralDin, Nisreen. (2012). "Estimating Victimization Rates. Trends and Risk Factors in Palestine". Unpublished M. Sc. Thesis; Department of Statistics. Faculty of Economics and Administrative Sciences. Al-Azhar University – Gaza. Gaza. Palestine.
- PCBS. (2008). Victimization Survey. Palestinian Central Bureau of Statistics; Ramallah. Palestine.
- Roodman, Allison. (2000). a Test of a Model of Sexual Victimization: A Latent Variable Path Analysis. Ph. D. Dissertation the Faculty of the Virginia Polytechnic Institute and State University.
- Sampson, M. (1985). Violent crime. Income inequality. and regional culture. Another Look. *Sociological Focus*. 18.199-208
- Stein, Rachel E. (2009). "A Comparative Analysis of Assault Victimization: The United States. The Netherlands. and England & Wales". *International Journal of Criminal Justice Sciences*. 4(1). 44-59.
- Straus, M. (1986). Domestic violence and homicide antecedents. *Bull N Y Acad Med.*;62:446-465

- Tyler, Kimberly A. Whitbeck, Les B. Hoyt, Dan R. & Cauce, Ana Mari. (2004). Risk Factors for Sexual Victimization among Male and Female Homeless and Runaway Youth. Sociology Department. Faculty Publications. Paper 35. Can be found in: <http://digitalcommons.unl.edu/sociologyfacpub/35>
- Von Hentig, Hans. (1948). “The Criminal and his Victim. Studies in the Sociobiology of Crime”. Psychoanal Q. Yale University Press. 19:105-106.