Repeat Property Criminal Victimization and Income Inequality In Brazil

José Raimundo Carvalho
CAEN/UFC and Laboratório de Econometria (CAEN/UFC), Brazil

Sylvia Cristina Lavor
Laboratório de Econometria (CAEN/UFC), Brazil

Abstract

This paper is concerned with the analysis of the impact that socioeconomic characteristics of individuals and regions, with special emphasis on income inequality, have in the process of repeat property criminal victimization. To accomplish this task, a series of count data models is estimated. The zero inflated negative binomial model performs best, in a series of nested and non-nested inferencial procedures. These econometrics models are applied to a unique data set: the 1988 PNAD. For this specific year only, an additional set of questions provide us with a national sample of victimization, something yet not available again until now. Our results are in close accordance to the international literature, and the estimative concerning the impact of income inequality on repeat victimization is quite strong and robust to the use of different measures and control for cluster effects. The elasticity of the GINI coefficient is positive and the strongest among all variables (0.3666), meaning that increasing inequality leads to more victimizations. The results about the effect of inequality has the potential to offer new insights for security policies at both the federal and state level in Brazil.

Keywords: Victimization, Inequality, Count Models

JEL Classification: K42, A13, C25

Resumo

Este artigo se propõe a analisar o impacto das características socioeconômicas do indivíduo e da região, com ênfase sobre a desigualdade de renda, no processo da vitimização criminal de propriedade repetida. Para atingir esse objetivo, modelos de dados de contagem são estimados. O modelo de zeros inflados negativo binomial apresentou melhores resultados, em uma série de procedimentos inferenciais aninhados e não aninhados. Esses modelos econométricos foram aplicados para uma base de dados singular: a PNAD de 1988. Para este ano específico, uma série de questões adicionais forneceu uma amostra nacional sobre a vitimização, algo até então não disponível. Nossos resultados se mostraram de acordo com a literatura internacional e a estimativa do impacto da desigualdade de renda sobre a vitimização foi bastante forte e robusta ao uso de diferentes medidas e controle de efeitos de *clusters*. A elasticidade do coeficiente da variável Gini foi positiva e a maior dentre todas as outras variáveis (0,3666), significando que um aumento na desigualdade leva a mais vitimização. Diante disso, os resultados obtidos acerca do efeito da desigualdade oferecem novas leituras para as políticas de segurança tanto em nível federal quanto estadual no Brasil.

1. Introduction

During the last years violence and crime has become one of the most prominent issue in the media and all over the population in general. More specifically, crime is seen as a restriction to the economic and social development of a region and, as a consequence, formulating effective policies to fight against crime appears to be a question with high social and economic returns (see Barslund et alii 2007).

Araújo e Fajnzylber (2001) consider that the phenomenon of criminality consists of a problem that presents diverse facets: it is social, because it affects directly the quality and the life expectancy of populations; it is economic, because it limits the potential development of economies; and it is political, since the necessary actions for fighting against crime involve the active participation of governments and the respective allocation of public resources. With significant increases in the rates of criminality and the growing importance given to the subject, governments started to face a challenge to formulate and to implement policies that can prevent and, consequently, reduce crime. To help the elaboration of these policies, it is of utmost importance the development of studies that allows the understanding of the determinants of crime activity.

Victimization, a different but closely related concept to crime, has not yet received the necessary full attention from economic scholars, however. Victimization has long been recognized by criminologists as an important perspective to understand crime. The appearance of studies on the subject of victimization made possible to analyze the socioeconomic characteristics of the individual and/or the family who suffered any type of delict and, also, to verify the determinants that influence the repeated victimization process. In other words, these studies had favored the identification of characteristics that contribute to most individual victimization. Some empirical studies had shown that is important to make research on victimization, mainly focusing on the reasons that conduce to its repeated incidence. According to Laycock (2001), it is possible to create and to implement programs to help fighting

⋆ Submitted in November 2009, accepted in July 2009.
E-mail address: josecarv@ufc.br

against recurrent victimization, that can be able to build a strategy for crime reduction. A key issue that received a great deal of attention is the impact of income inequality on victimization.

Some studies, such as Fajnzylber e et al. (1998), make use of inequality income indicators to show that the greater the concentration of income in a determined place, the greater will be the incidence of delicts. Also, papers as Levitt (1999), using data on the average income of the family, indicate that, in contrast to what people might imagine, poorer people have greater chances of suffering with victimization than the richer people do. Although there exists a great deal of discussion about this subject, there are few papers that study victimization at the individual level, in other words, that take in consideration the role played by the victim at the moment of the delict. The majority of studies analyzes crime under the perspective of the delinquent. However, it is necessary to make clear that there exist two ways to study crime: focusing on the side of the criminal, studying the reasons why he/she has committed the crime; or analyzing the side of the victim, verifying which characteristics he/she possess that can influence his/her own victimization.

In face of this fact, this paper intends to analyze, in a more inclusive form, the victimization process in Brazil, giving emphasis to the characteristics of repeat victimization. In other words, a brief analysis of the characteristics of the individual that became a victim is made, independent of the amount of times that the event manifested. And also, the determinants of the process of repeat victimization. At the same time, an analysis on the relation between inequality of income and victimization in the country is performed, addressing the question whether income concentration affects the incidence of crimes and the behavior of the victims. For the accomplishment of this task, the data of the National Research for Sample of Domiciles (PNAD) (see Brazil 1988), a unique occasion that a data collection with a national character approached questions related to the subject of the victimization, is used. Although it is not officially a survey on victimization, the data set made possible an analysis regarding victimization in Brazil. To complement this database, information on economic variables through the Institute of Applied Economic Research (see Brazil 2008) for the same year, had been used.

By means of a sequence of count data econometric models, we are able to replicate international results, as well as get new ones, concerning the positive effects [1] of variables such as urban local of residence, being the head of the household and status of employed. Sex, surprisingly, has no statistically significant effects. The guardianship effect, where households with a larger number of components have lower mean number of victimizations, is found in our data set. The effect on income is positive, prompting us to conclude

[1] Positive effect means that an increase in the independent variable leads to an increase in the mean number of repeat victimizations.

that the effect of attractiveness of the target dominates the effect of private protection.

The key issue under analysis, i.e., the effect of income inequality, proxied by the Gini index, on repeat victimization bring us surprising new evidence. The effect is positive and turns to be the largest among all independent variables! Actually, this result is robust to the use of different measures such as the Theil index, as well as to correcting for cluster effects and heterocedasticity. As far as we know, our paper is the first to obtain this result using a methodology based on count data models that brings together micro as well as aggregate variables as explanatory factors. Even though our results must be taken as not definitive, they point out to some important issues that must be considered in the formulation of policies of development and security so debated among political and academic spheres in Brazil.

The structure of the paper is comprised, besides this introductory section, of more 5 sections. Section 2 makes a literature review about repeat victimization and income inequality. Section 3 performs a detailed analysis of the available empirical evidence on repeat victimization, using the 1988 PNAD as the main source. Section 4, in a rather detailed way, develops the main count data models that are used for estimation. Section 5 presents the estimated parameters as well as perform the necessary inferential procedures. Section 6 concludes the paper, offering some avenues for future research.

2. Literature Review

The majority of studies that currently exists in the literature on crime had its initial motivation in Becker (1968), who shown that the relation between economic incentives and criminality is a consequence of rational behavior. In other words, Becker suggested that the criminal actions can directly be seen as connected to the rational decisions of the criminal, where he/she observes the costs/benefits of the crime commitment.

Although Becker's original work prompted a huge amount of research, with many other empirical studies connected to the assumption of rational choice, the criminological literature has pointed out the importance of paying attention to the victimization process. The main point about the victimization process is that a crime does not necessarily turns out to be recorded as a crime. In fact, as it is well known to criminologist and sociologist but unfortunately less well known among economists, crime under-reporting is far from being just a curiosity with no systematic explanation. The "dark figure", i.e., victimizations that are not officially reported, is a well established regularity. Among the very few economic papers that recognized this important fact are MacDonald (2001) and MacDonald (2002). For Brazil, it is worth mentioning the paper of Santos e Kassouf (2008).

Before we go into details on some papers, it is important to digress a little bit about the concept of repeat victimization. Given a specific time range, the larger this period the higher the probability of one experiencing one or more episodes of victimization. This issue was incorporated, with great success into the criminological main stream by Pease (1998). By noting that the best predictor of victimization is past victimization, the author developed an approach of criminal policy based on the notion of repeat victimization. In fact, now is well known that the amount of victimization is heavily concentrated among few people (see, Pease (1998), Farrell e Pease (2001) and Tseloni e Pease (2003), among others).

The literature on repeat victimization is now huge,[2] however is almost entirely located among criminologists and few sociologists. Repeat victimization measures are entirely neglected by econometricians and economists. Even though discussing the perils of time aggregation of the number of victimizations under a single indicator of victimization occurrence is beyond the scope of our paper, it is important to stress some possible bias of this approach. First, there is the problem of completely missing the fact that victimization is concentrated among certain types of individuals. Second, the number of victimizations in a time interval carries along many behavioral insights. Finally, count models can be rationalized by assuming the existence of heterogeneous groups of people: one that never experience a victimization and the other that are comprised of people who can experience any amount of victimization, including zero. By aggregating repeat victimization we loose the appeal of applying such count data models. So, we are going to treat the victimization process as one of repeat victimization.

As mentioned before, our approach take a different view by using victimization as the key dependent variable in a model of crime. This is a growing tendency among economists and econometricians. So in order to study the impact of income inequality on crime, we make a point in favor of using (repeat) victimization as the best choice. We are not alone in this regard. The paper of Gaviria e Pagés (1999) analyzed the relation between victimization and income classes. They employed a data set from 17 countries from Latin America for the years 1996 and 1998. They find that the majority of property theft come from the high and middle classes who live in larger cities. Besides that, city growth is correlated with victimization rate. However, they use the Latinobarometer data set (Latinobarometer 1999) which falls short as being a good victimization survey.

For a more geographically focused study, it is worth mentioning Tella et alii (2002). These authors try to asses the correlation between income distribution and victimization. The key research question is to figure out which social group suffered the most (in terms of higher victimization rates) with the growing rate of criminality in Argentina. They used a survey questionnaire to collect information on victimization at metropolitan Buenos Aires. For household

[2] For a recent paper and account of the issue, see Tseloni e Pease (2004).

victimization, poor people suffered more with increasing criminality. They try to explain this fact by arguing that richer people are able to invest in private security, something less likely for poor people. Also, the very private spending on private security from the part of rich people creates negative externalities [3] for poor people. Although their results are robust, they do not consider repeat victimization.

Another important paper is Bourguignon et alii (2003). Even though they do not use victimization, their careful analysis of the impact of income inequality and crime is illuminating. The authors develops a "structural" model of crime. Their main point is to rationalize some negative results regarding the effect of inequality on crime. They argue that usual measures of inequality, such as Gini and Theil indexes, fail to represent the inequality change that really matters for criminal activity: those changes located in a specific income range, close to the bottom of the distribution. After developing a clever model, they estimate it with Colombian data and find the "right" effect, i.e., the higher the inequality the greater the crime rate. Note that, as we will argue later on, it might be possible to rationalize their result by saying that crime should be measured by victimization. It means that it is likely that some existing "puzzles" in the crime literature might be due to the economists lack of consideration about the difference between crime and victimization.

Finally, Barslund et alii (2007) employ a micro data set from Mozambique to study the determinants of victimization. Among other correlates, they include income. Their main results point out to a robust positive effect of income on victimization, although at decreasing rates. Inequality, as well as unemployment have the effect of increasing victimization rates. Interesting, their results show that even though the poor is less likely to suffer thefts [4] they suffer more, relatively to rich people, as they have a lower chance of recomposing the lost what has been taken. Despite the revealing results, they restrict themselves to victimization, disregarding the more important issue of repeat victimization.

Another stream of research related to ours started with Levitt (1999). This seminal paper initiated an approach of studing victimization that has as the main focus a preoccupation with questions of crime distribution and inequality in income distribution. Researchers who study this issue have found constantly that crimes are distributed differently among social groups. [5] Therefore, the central question to be examined consists of whether an increase in inequality of a region' income impacts in the victimization and which social groups suffer more with this.

[3] By displacing criminals to poor areas.

[4] Of course, this is a reasonable fact.

[5] Note that such approach can be found in the initials writings of Becker. In fact, in his work it has been recognized that a rich person becomes a more attractive target when her income increases. However, if the rich behaves in such a way that it reduces his risk of victimization, by investing in private security, then his risk of victimization can actually diminishes. Consequently, the poor' victimization can increase in relation to the rich one.

Using average income as inequality indicator, Levitt (1999) asserts that due to the income inequality increase in the United States, the less-favored people have started to suffer more with victimization, because crimes started to be more concentrated in the poor areas. A possible explanation for the reduction of the victimization risk among rich people, according to the author, could be the increase of private security expenses. In other words, with greater purchasing power, one can invest more in this type of security. Beyond investigating victimization among social groups, Levitt (1999) also analyzes this phenomenon along the race dimension. According to the author, not whites who has a determined income are victimized more frequently than whites. Moreover, the former are twice more vulnerable to be robbery victims than the later; and those who are not white or rich, get twice as more chances to have their vehicle stolen than the rich whites.

An important question about the difference in the victimization distribution among social groups is that income level influences the way that a richer person protect herself from some delict. As discussed previously, well-taken care with private security are recognized by empirical studies as capable to reduce the level of crime in some areas, but can increase in others, because this type of protection can generate negative externalities in poor areas.

Another paper that approaches the relation between income inequality and victimization is Thacher (2004). The author shows that despite the decline in criminality during the period from 1974 to 2000, in the United States, this benefit was not shared equally. Said in other words, poor people have started to suffer more with victimization than people with higher income in all the types of delict, confirming the results gotten in other studies, in which increasing the inequality of income distribution, impacts more individuals with lower purchase power in terms of victimization.

In Brazil, most of the debates on this issue refer to the influence of socioeconomic factors on criminality, beyond the measurement of the cost of violence for society. However, the available studies that investigate the problematic around criminality are limited. Among these few, Resende (2007) verifies the relation between income inequality and criminality. Regarding studies about the link between victimization and inequality we found only one paper in Brazil. Gomes e Paz (2007) builds a "structural" model of victimization that serves as a basis for the estimation of a binomial model of victimization for the population of São Paulo, the largest Brazilian state. Even though their main focus is on the city size effect on victimization, they are able to find an increasing non-monotonic income effect on victimization. However, they do not address inequality in income issues, nor they include an analysis of multiple victimization.

In face of this literature review, we believe we identified some important gaps, both nationally and internationally, in the literature of victimization and income inequality that are worth pursuing, such as:

(i) Very few international studies employ repeat victimization measures and no national study does it;

(ii) The data set used are generally comprised of small samples;[6]

(iii) None paper in Brazil included inequality measures as a possible predictor of repeat victimization.

Hence, in face of these gaps, this work will focus, from now, on the repeat victimization process in Brazil, verifying which individual characteristics can impact on victimization occurrence and whether income inequality influences this phenomenon. The data set used, as mentioned, is the 1988 PNAD.[7] The econometric model is based on the estimation of count data regressions that will try to control for both unobserved heterogeneity and excess of zeros present in the sample. Next section will address the empirical evidence.

3. Data

Recently, some empirical studies have shown great interest in repeat victimization issues. In the last two decades, it was recognized that repeat victimization of people and/or places represents the majority of all occurred victimization in some regions, e.g., Kleemans (2001). Moreover, Sparks (1981) showed that this issue occurs more frequently than it could foresee.

In accordance with Ybarra e Lohr (2002), there are many evidences that victimized people have a higher probability of suffering future victimizations than those ones that had never been victims. Such evidence can be observed by following individual behavior after the occurrence of a delict. Moreover, the study of Kleemans (2001) reveals that repeated victimization is more likely to occur in areas with high criminality rates, suggesting that crime concentration is linked with the number of victimizations.

In face of this evidence, it can be said that the odds of a person becoming a victim, considering that he/she has already suffered some type of delict, are increasing. Said this, this section analyzes the available empirical evidence using as the main source of information on victimization the 1988 PNAD. This research is the most important household survey employed in Brazil. Besides the data contained in PNAD, we will use other empirical evidence about the economic characteristics of the region, for instance, Gini index of each state, from the IPEADATA, for the same year and other measures of inequality and poverty. Even though that data set is not a victimization oriented survey, in

[6] With the exception of American and British studies that use the Criminal Victimization Survey and the British Victimization Survey, respectively. These data sets are large and well designed surveys specifically developed to track the pattern of (repeat) victimization.

[7] Admittedly this is a rather old data set! So, any results coming out of this analysis should be interpreted as only illustrative. However, if one considers both the richness of the sample, its national coverage as well as the fact that there is no national victimization survey in Brazil as to July, 2009, the 1988 PNAD has a lot of appeal. As a matter of fact, the Brazilian government is planning its first national victimization survey for the second semester of 2009.

1988 it brought as a special feature a small set of questions amenable to be used in a victimization study. The exact wording of the questions were:

(i) "During the period of October/1987 – September/1988 have you ever been a victim of attempted burglar or attempted theft?"

(ii) "During the period of October/1987 – September/1988 have you ever been a victim of burglar or theft?"

For the sake of simplicity and because we lack a theory for attempted victimization, we aggregate attempts and actual victimizations under the common denomination of victimization and show its distribution on Table 1. In accordance with that table, the distribution is overdispersed, with variance higher than the mean (Mean=0.12 and Var=0.32). Moreover, this sample is composed of a high proportion of zeros: almost 92%. So, negative binomial and zero inflated models seems to be the most appropriate to model repeated victimization. The explanation for this choice will be better displayed in the section of Methodology.

Table 1
Frequency of victimizations

Count	**Victim (%)**
$P(Y = 0)$	91.75
$P(Y = 1)$	5.94
$P(Y = 2)$	1.49
$P(Y = 3)$	0.41
$P(Y = 4)$	0.20
$P(Y = 5)$	0.08
$P(Y = 6)$	0.06
$P(Y = 7)$	0.01
$P(Y = 8)$	0.02
$P(Y = 9)$	0.00
$P(Y \geq 10)$	0.04
Total	100.00

Source: Elaborated by the authors.

Another feature of victimization data sets is the fact that the conditional distribution of having two or more episodes of victimization is increasing (at least from the first values of victimizations). This implies that the greater past victimization experiences the person has suffered, the higher the likelihood of becoming a victim again. This result coincides with the same ones gotten from Ybarra e Lohr (2002) and Kleemans (2001). Table 2 shows this pattern.

Table 2
Conditional victimization

Probability	**Result (%)**
$P(Y = 1)$	5.94
$P(Y = 2 \mid Y \geq 1)$	18.00
$P(Y = 3 \mid Y \geq 2)$	17.63
$P(Y = 4 \mid Y \geq 3)$	24.34
$P(Y = 5 \mid Y \geq 4)$	19.28
$P(Y = 6 \mid Y \geq 5)$	27.65
$P(Y = 7 \mid Y \geq 6)$	9.33
$P(Y = 8 \mid Y \geq 7)$	20.63
$P(Y = 9 \mid Y \geq 8)$	5.71
$P(Y \geq 10 \mid Y \geq 9)$	92.41

Source: Elaborated by the authors.

For our purposes, questions related to personal victimization will be analyzed, this means that will be given an approach to the question of whether the individual was a victim of robbery or theft during the period from October of 1987 to September of 1988, and how many times this occurred. Hence, the model to be estimated possesses as dependent variable the amount of times that an individual was victim of robbery or theft or suffered an attempt of robbery or theft, assuming discrete values bigger or equal to zero. Moreover, the model has as independent variables the socioeconomic characteristics of the individual, such as where he lives, sex, color, if he is or not the family head, if he worked in the same week that he was interviewed, age, age squared, scholarly, number of components of his family, per capita familiar monthly income and information on state income inequality, proxied by the index of Gini.

Table 3 briefly describes the variables used in the model and shows the main descriptive statistics. By inspection of Table 3, it is already possible to make a preliminar analysis on the variables. Accordingly, great part of the sample is found in the urban zone. Moreover, a little more than 50% of the sample is formed by women and people who considered themselves to be of white color. More than half of the sample consists of people who said that they are not heads of the family in the household where they live and that they work in some activity. Another point to be detached is that the average age of the interviewed people is 32.8 years and that the average time of studies for them is almost five years. A percentage of 81 is married and 75 % of households own a TV equipment, and 81 % of all household's heads is married (variable CIVIL).

As to the inequality characteristics, Gini index is, on average, 0.59. Next section

Table 3
Description of the dependent and independent variable

Variable	**Description**	**Min**	**Max**	**Mean**	**Var**
Y	Counts of victimization	0	40	0.12	0.32
SIT	Dummy for localization (Rural = 0 and Urban = 1)	0	1	0.79	0.164
SEX	Dummy for Sex (Female = 0 and Male = 1)	0	1	0.48	0.250
RACE WHITE	Dummy for race (Non-white = 0 and White = 1)	0	1	0.52	0.249
CONDD	Dummy for Condition in the household (Non head = 0 and Head = 1)	0	1	0.35	0.227
AGE	Age in years	16	60	32.8	146.89
AGE2	Age squared	256	3600	1222.88	773555.45
ACTIVITY	Dummy for employment status (Unemployed or out of the labor force = 0 and Employed = 1)	0	1	0.66	0.225
EDUCATION	Years of education	0	10	4.94	11.099
NUMCOMP	Family size	1	28	5.20	6.027
LNINCOME	Log of per capita familiar income (in 1,000 of Cz$)	1.61	20.72	11.36	2.793
GINI	Gini index	0.4812	0.6573	0.5945	0.001
TV	Own a television equipment	0	1	0.75	0.189
CIVIL	married = 1; single =0	0	1	0.81	0.154

Source: Elaborated by the authors.

develops a number of traditional count models. Also, a more detailed discussion of the economic and econometric issues that arise by using the GINI as a right hand side variable, a variable that stays constant at the state level, is developed.

4. Econometric Methodology

Models of counting data are appropriate to estimate number of occurrences of events. In this case, the dependent variable Y assumes non-negative integers $(0, 1, 2, \ldots)$, that correspond to the number of events realized in a determined period of time. For instance, events should be: number of visits to a doctor that a person makes during the month; number of times that someone was assaulted in the last year, number of purchased items in a supermarket and so on. Hence, for our analysis, i.e., repeat victimization, count models should be the most appropriate approach. Before developing the econometric model it is worth digressing a little bit about the different types of count models.

A straightforward, but usually incorrect approach, is to apply linear regression directly. Even though there exists situations in which the linear regression model offers reasonable results, is much safer to use models specifically designed to deal with count data. Some models under that label are: Poisson Regression Model (PRM), Negative Binomial Regression (NBR) and variations of these models, as the Zero Inflated Poisson (ZIP) and the Zero Inflated Negative Binomial (ZINB) regression. Generally, the PRM model is the starting point for analysis, even though its use could be so frequently inadequate (see Cameron e Trivedi 2005).

Let Y be a variable that indicates the number of times that an event occurs along a specified time interval. If Y has a Poisson distribution, then:

$$Pr(Y|\mu) = \frac{e^{-\mu}\mu^{Y}}{Y!} \quad for \quad Y = 0, 1, ... \tag{1}$$

A key fact about this model is the property that $E(Y) = Var(Y)$. Note also that $\mu > 0$ is the only parameter defined in the distribution. The Poisson regression model (PRM) is derived from the distribution of Poisson for the parameterization of the relationship between the mean μ and a vector of regressors X. The standard hypothesis is to use the following structural equation:

$$\mu_i = E(Y_i|X_i) = \exp(X_i'\beta) \tag{2}$$

Again, the mean is equal to the variance, i.e., $\exp(X_i'\beta) = Var(Y_i|X_i) = \mu_i$, meaning that the Poisson regression is intrinsically heteroskedastic.

However, our sample data presents different values for mean and variance, which goes against the property of the Poisson distribution. Such kind of pattern is typical of data sets on repeat victimization. So, following the proposal of Tseloni and Pease (2003), we need to find a better specification to deal with overdispersion. The negative binomial regression model (NBR) corrects the imperfection of the PRM by adding a parameter δ_i that reflects the heterogeneity unobserved between the observations. The NBR adds an error ε that is assumed to be uncorrelated with de components of the vector X.

$$\widetilde{\mu_i} = \exp(X_i'\beta)\delta_i \tag{3}$$

where $\delta_i = \exp(\varepsilon_i)$. To identify the model, it is assumed that $E(\delta_i) = 1$, which corresponds to $E(\varepsilon_i) = 0$ in the linear regression model. It is straightforward to see that now we have $E(\widetilde{\mu_i}) = \mu_i E(\delta_i) = \mu_i$, i.e., the NBR model has the same mean as the Poisson regression model, however its variance allows for over-dispersion. We model overdispersion due to the unobserved heterogeneity between the individuals. The probability distribution for the NBR model must be conditioned on both the values of X as well as on the value of δ_i:

$$Pr(Y_i|X_i, \delta_i) = \frac{\exp^{-\widetilde{\mu_i}} \widetilde{\mu_i}^{Y_i}}{Y_i!} \tag{4}$$

Since δ_i is unknown, one cannot calculate $Pr(Y|X)$. However, this could be overcome by assuming that the uncertainty about δ_i can be represented by a gamma distribution. Then, $Pr(Y|X)$ can be calculated as a weighed combination of $Pr(Y|X, \delta_i)$ for all the values of δ_i, where the weights are determined by $Pr(\delta_i)$. So, the binomial negative distribution is given by:

$$Pr(Y|x) = \frac{\Gamma(Y - \alpha^{-1})}{Y!\Gamma(\alpha^{-1})}(\frac{\alpha^{-1}}{\alpha^{-1} + \mu})^{\alpha^{-1}}(\frac{\mu}{\alpha^{-1} + \mu})^{Y} \tag{5}$$

where Γ is a function gamma. Despite of the frequent use of these models, it is common to find counting data sets presenting excess of zeros.[8] In this situation, the application of these models can generate inconsistent estimates.

For zero inflated models, the dependent variable Y_i assumes again non-negative integer values. However, there is a crucial difference in those models. There are two unobserved sub-populations: those who will never experience any count and those who could experience any number of counts, including zero. Let us define the first sub-population by $A = 0$ and the other sub-population by $A = 1$. Since by assumption group membership is not observed, one can model group membership by, say, a logit equation:

$$Pr(A_i = 0|X_i) = F(X_i'\theta) \tag{6}$$

where F is the logit or probit function. For those belonging to the sub-population where $A = 1$ it is assumed that conditional on the event that $A = 1$, the probability distribution of the number of counts follows a specific count model. For instance, if we assume a Poisson model, the outcome will be the Zero Inflated Poisson Model (ZIP), otherwise if we assume a negative binomial model, it follows the Zero Inflated Negative Binomial Model (ZINB). Note that technically the zero inflated models need to solve a mixing problem: there are two sub-groups, one with a specific marginal probability, i.e., $Pr(A_i =$

[8] In the repeat victimization literature, the observed proportion of zeros is frequently beyond 85 %.

$0|X_i)$, and the conditional distribution of counts $Pr(Y_i = j|X_i, A_i)$, for $j = 0, 1, 2, \ldots$. The observed part of the model comes from the following mixing operation:

$$\begin{aligned} Pr(Y_i = j|X_i) =& Pr(Y_i = j|X_i, A_i = 0)Pr(A_i = 0|X_i) \\ &+ Pr(Y_i = j|X_i, A_i = 1)Pr(A_i = 1|X_i) \end{aligned} \tag{7}$$

The probability of the excess of zeros is denoted by $\pi_i = Pr(A_i = 0|X_i)$, where $0 \leq \pi_i \leq 1$. The mean and the variance of the ZIP random variable is given by:

$$E(Y_i|X_i) = (1 - \pi_i)\lambda_i \tag{8}$$

$$Var(Y_i|X_i) = (1 - \pi_i)\lambda_i(1 + \pi_i\lambda_i) \tag{9}$$

If Y_i followed the ZINB distribution, then the mean and variance are given by:

$$E(Y_i|X_i) = (1 - \pi_i)\lambda_i \tag{10}$$

$$Var(Y_i|X_i) = (1 - \pi_i)\lambda_i(1 + (k + \pi_i)\lambda_i) \tag{11}$$

where k is a parameter that indicates the over-dispersion. The ZINB model reduces itself to the ZIP one when $k \to 0$.

We turn now to the less straightforward question regarding how to justify the use of an aggregate measure in the right hand side together with an individual-level dependent variable, i.e., the number of victimizations. A key issue that must be explained is the possible implications of using that aggregate measure of inequality, the GINI coefficient. In fact, there are two different, but interrelated, issues that arise: the first point has to do with how economists incorporate neighborhoods or contextual effects, and the second issue has to do with how econometricians handle such situation.

In fact, economists have been specially reticent in using neighborhood effects (or contextual, or social effects) in their models, even though papers like Dickens e Katz (1987) and Topel (1986) have employed freely such approach, to name just a few. This is in contrast to sociologists, criminologists and epidemiologists (see, Blalock (1984), Miethe e McDowall (1993) and Wen et alii (2003), respectively). The use of contextual (or, ecological) effects is commonly accepted as an important theoretical benchmark by those scholars. However, in accordance to Durlauf (2004), this trend is changing and economists are increasingly recognizing that there are other sources of effects besides only those mediated by the price mechanism or budget constraints. Durlauf (2004) provides many examples of neighborhood effects that have been explicitly justified along sociological and/or psychological lines. However, from the onset we would like to stress that we have opted to refrain from being characterized as a paper belonging to the neighborhood effects tradition. In fact, by using the GINI coefficient we employ a sort of "aggregate effect", a more preferable denomination, since the label neighborhood effect is more appropriate to levels of aggregation lower than a state. Anyway, we offer below how such "aggregate effect" can be justified for the case of individual victimization.

There is a traditional explanation for why aggregate inequality affects crime (see specifically, Bourguignon et alii 2003). Other things being equal, the expected gain from crime (especially property crime) depends on the mean income of the population of potential victims, whereas the cost of crime commitment depends, by means of the opportunity cost of time and punishment in case of detection, on the income of the potential criminal. From such logic, and invoking another *ceteris paribus* argument, the risk of victimization depends on how the income is distributed among the population, since the process of victimization is intrinsically related to the process of criminal activity. Actually there are many papers by economists evaluating the impact of inequality on crime (for instances, Bourguignon et alii (2003), Fajnzylber et alii (2002a) and Choe (2008), just to name a few), but as far as we know none use an approach analogous to ours and none applies their modelling strategy to model the repeat victimization process.

The second issue is completely econometric in nature and was recognized by the seminal papers of Moulton (1986) and Moulton (1990). The essence of Moulton's effect is the impossibility of keeping the assumption of independent disturbances when using aggregate right hand side variables in models where the dependent variable is less aggregated, or in another words, when employing aggregate variables to explain micro units behavior [9]. In fact, Moulton was able to show striking results concerning the effect of not controlling for the correlation errors within groups. Many aggregate variables that shown a strong effect (in the sense of being large and statistically significant), after considering the correlation within groups turn to be weaker and many times not statistically significant.

Using the modern parlance, Moulton's problem is an issue dealt with cluster sample models. Accordingly to Wooldridge (2006), the econometric literature on how to deal with cluster sample is still evolving, but there are some developed approaches to deal with such problems, specially in a linear context. In a non-linear context, our case, the field is still open. However, for those data set where the number of clusters is small in relation to the number of observations per cluster there are some remedies provided by commercial software. So, since our aggregate variable is constant within the state, and we have a great number of observations per state, we are going to employ an estimation approach that

[9] However, Moulton's effect might have far reaching implications, since it appears that econometricians left completely unquestionable the use of aggregate dummies (say, state, region dummies) in microeconometric models!

considers that fact in our count data models estimation. The next section provides the estimatives from the count models.

5. Empirical Analysis and Results

In order to address the issue of the impact of economic inequality on repeat victimization we estimated four models: the Poisson model, the Negative Binomial Model, the Zero Inflated Poisson Model and the Zero Inflated Negative Binomial Model. However, given the inferential procedures applied to the data we opted to showed only the results.[10] for the NBRM and ZINB. From Table 4 we are going to interpret the estimated elasticities (right columns) and try to justify the choice of the "best" model.[11]

First, note that all models are estimated taking in consideration the effect of the cluster imposed by using the GINI variable as an aggregate effect. In fact, the use of an aggregate independent variable act "as if" the sample was collected by means of a clutering scheme (see Wooldridge 2002, 2006). Hence, all estimation results are performed using the option "cluster" of the software STATA version 10 which combines the cluster effect with robust standard errors. By inspection of Table 4, it was verified that all estimated elasticities are statistically significant at 10%, with the exception of SEX, RACE WHITE and GINI. In this sense, it appears that these variables do not influence the occurrence of repeat victimization. As a first specification test, we reject the PRM in favor of the NBRM as the value for the test of $H_0 = \alpha = 0$ promptly rejects this null ($LR\ \ \chi^2(1) = 41000$). Next, we test the ZIP model against the ZIBN model by means of a likelihood ratio test. The ZIP model is reject[12] in favor of the ZINB model. Hence, we are left with the two models presented on Table 4.

These models are non-nested, though. In order to get a first insight regarding model specification correctness, we plot the mean predicted probabilities for the NBRM and ZINB. Figure 1 plots the difference between observed probability and mean predicted probability for both model at different values of Y. The ZIP model performs much better for $Y = 0, 1, 2$ and 3, however, the ZINB outperform the ZIP models for higher values of Y. So, at best we reached an inconclusive result by the visual inspection of the graph. Any more definitive result regarding the choice between these two models must be based on a non-nested test. We opted for the widespread Vuong Test. The value of this statistic appears in Table 4. The value of $z = 30.36$ flatly rejects the NBRM in favor of the ZINB. Hence, our comments on the estimated elasticities are based on our remaining "best model", i.e., the Zero Inflated Negative Binomial

[10] In fact, some models are nested, and others could be tested by means of non-nested procedures.
[11] Estimated values presented with their robust standard errors between parenthesis. An "***" means significant at the 10% level.
[12] This results are omitted but could be obtained from the authors.

Model, even though the results for the NBRM are presented together. Next, we comment on the effects of the independent variables.[13]

Before we start commenting on the elasticities values, it is worth remembering a key structural interpretation of zero inflated count models. From inspection of Equation 6 it is clear that any zero inflated model is a mixing of two type of populations: those who will never experience any count and those who could experience any number of counts, including zero. This means that variables appearing in the left column of the **Zero Inflated Neg. Binomial** model of Table 4 impact on the expected count for those belonging to the population that could experience any number of counts, including zero. Variables appearing in the right column (*Inflate*) correspond to the binary model predicting group membership. So, a possible "structural interpretation" goes as the following: there are some variables that could "shield" people from being victims, in contrast to other variables that, although modifying the risk of being victimized, can not guarantee complete protection. The variables SEX, RACE, WHITE and TV belong to the latter group and all other variables belong to the former.

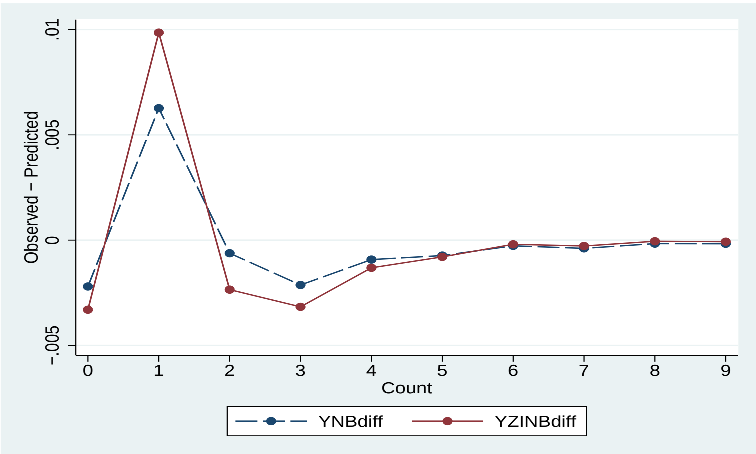

Fig. 1. Mean predicted probabilities

All estimated elasticities are significant at a 10 % level, with exception of SEX.[14] Living in an urban area increases the number of victimizations,

[13] In order to follow the discussion the reader must track the estimated elasticities values for the zero inflated negative binomial model. The interpretation for elasticities is straightforward: they give the percentage change on the mean expected number of victimizations given a small percentage change on the independent variable. These values are located at the last column with robust standard errors values under parenthesis, at the bottom.

[14] We have used some other suggestions for independent variables such as state and regional dummies, past registration of victimization, circunstantes of past victimization occurrence, offered by the referee. However, they turned out to be discarded as non significant during the estimation process. Also, another aggregate variable used in an earlier version of the paper, i.e., POOR (the percentage of poor households in the state) has no explanatory power when used in conjunction

Table 4
Estimation results

	Negative Binomial		**Zero Inflated Neg. Binomial**		
Var.	**Coef.**	**Elast.**	**Coef.**		**Elast.**
				Inflate	
SIT	0.5474	0.0508***		-1.1759	0.0788***
	(0.0938)	(0.0082)		(0.2048)	(0.0098)
SEX	-0.0306	-0.0033	-0.1056		-0.0140
	(0.0469)	(0.0051)	(0.0772)		(0.0101)
RACE WHITE	0.1620	0.0174	0.1580		0.0209***
	(0.0949)	(0.0113)	(0.0845)		(0.0119)
CONDD	0.7372	0.0908***		-1.9826	0.1351***
	(0.0864)	(0.0096)		(0.2873)	(0.0091)
AGE	0.0457	0.0049***		-0.0750	0.0054***
	(0.0129)	(0.0013)		(0.0110)	(0.0007)
AGE2	-0.0004	-0.0001***		0.0008	-0.00006***
	(0.0001)	(0.0000)		(0.0001)	(0.00001)
ACTIVITY	0.2281	0.0237***		-0.3656	0.0264***
	(0.0391)	(0.0039)		(0.0533)	(0.0042)
EDUCATION	0.0944	0.0101***		-0.1887	0.0137***
	(0.0074)	(0.0012)		(0.0132)	(0.0015)
NUMCOMP	-0.0325	-0.0035***		0.0791	-0.0057***
	(0.0157)	(0.0015)		(0.0140)	(0.0011)
LNINCOME	0.0589	0.0063***		-0.1406	0.0102***
	(0.0262)	(0.0027)		(0.0281)	(0.0018)
GINI	3.2398	0.3491		-5.0184	0.3666***
	(2.2279)	(0.2330)		(2.5638)	(0.1787)
TV	0.1780	0.0183***	0.1867		0.2373***
	(0.0436)	(0.0048)	(0.0472)		(0.0060)
CIVIL	-0.2712	-0.0319***		0.4118	-0.0302***
	(0.1267)	(0.0162)		(0.0614)	(0.0052)
CONSTANT	-5.5209***		-1.3916***	8.2580	
	(0.7571)		(0.1336)	(1.6130)	
$\ln(\alpha)$	0.3359		1.4645		
	(0.1290)		(0.1035)		
α	1.3992		4.3256		
	(1.0213)		(0.4480)		
Log Likelihood = -58138.324			Log Likelihood = -57476.19		
$LR\ \ \chi^2(13) = 2917.10$			$LR\ \ \chi^2(3) = 17.64\ Prob > \chi^2 = 0.0005$		
$Prob > \chi^2 = 0.0000$			Voung test of ZINB vs. NB:		
Number of obs $= 162,686$			$z = 30.36\ \ Prob > z = 0.0000$		

Source: Elaborated by the authors.

although inelastically (0.0788). Heads of the household suffer more victimizations than non-head does (0.1351), probably due to the fact that they must expose themselves more to guarantee the survival of the household, i.e., they must work more hours, perform duties related to the household, go out

with GINI.

more often, carry more valuable objects and so on. The explanation for the effect of the variable ACTIVITY goes analogously. AGE has a positive effect and AGE2 a negative, but very small, effect. So, this parabolic pattern of victimization could be possibly be linked to the age profile of routine activities of individuals.

The level of education seems to have a positive effect on the mean number of victimizations. That results, apparently at odds, reveals an interesting fact. If we assume that the probability of being a victim is function of the attractiveness (i.e., likely amount of income that could be taken) of the target and of the level of private protection (represented by technologies of protection, knowledge and behaviors) the positive elasticity seems to indicate that the prevailing net effect is the target attractiveness. This means that even though more educated people know and do protect themselves better than the less educated ones, they appear to be better targets to the criminal's eyes. Family size, proxied by the number of household occupants (NUMCOMP), has a negative effect on victimization. Such effect, known as guardian effect, seems reasonable as larger families could protect their members better, just by acting as guardians. Per capita family income (LNINCOME) has a positive elasticity (0.0102). Being married decreases the risk of more victimizations and owing a TV increases the risk of victimization. Finally, being a married head of a household decreases the mean number of victimizations. So far, with few exceptions, our results confirm those obtained in the international literature, for instances, those obtained in Tseloni e Pease (2003).

The question regarding inequality of income and crime first appeared in a systematic approach on Ehrlich (1973). Ehrlich showed, both theoretically as well as empirically, a positive correlation between crime and income inequality. More recently Fajnzylber et alii (2002b) reach the same conclusion by means of a cross-country empirical analysis. However, as far as we know there is no past attempt to measure the impact of inequality on repeat victimization. We find a strong positive effect of inequality, measured by the Gini index.[15] In fact, the estimated parameter for the GINI variable, is positive, although inelastic, 0.3666! Roughly, a 1% increase in inequality generates approximately an increase of 0.36% in the mean number of victimizations. It is worth mentioning that the first estimated models, without correction for cluster effects, showed a stronger effect for the GINI variable coupled with smaller standard errors. After controlling for cluster effects, the elasticity for the GINI variable decrease considerably and the robust standard errors are considerably greater compared to the original (non-robust) standard errors, even though the coefficient remains statistically significant.

This evidence adds to the list of deleterious effects of income inequality and extends the scope of public policy. By establishing a link between inequality and victimization, governments might think more carefully about policies focused

[15] We found similar results using the Theil index. The results could be requested from the authors.

on crime prevention (victimization). The traditional view that crime must be fight with police expenditure must be compared to the returns of spending in policies targeted at decreasing income inequality. Next section concludes and offers some avenues of future research.

6. Final Considerations

The objective of the paper was to analyze the socioeconomic characteristics of individuals, also considering the income inequality indicators, and to verify how much they can influence the process of repeat victimization. Using data from the 1988 PNAD and the IPEADATA for the same year, it was possible to create an econometric model for counting data, to apply it to the repeat victimization phenomenon and estimate it. The results of this investigation appears to be compatible with the international literature on the subject.

The characteristics of individuals was shown to be important to repeat victimization. Living in urban areas turns people more vulnerable to suffer robbery or thievery crimes. This result was expected, since great part of crimes occurs in such areas. Moreover, being male, family head and to have some type of job also increase the person's probability to be victimized.

Years of studies, that, in accordance with economic literature, possess positive relation with the person income, also have a positive impact on repeat victimization. We offer two explanations for such counter-intuitive effect: either more educated people have a daily routine that expose themselves more often, or they signal more income.

It is not difficult to understand this, since more time of schooling results in person having more qualifications and more opportunities, this in turn increases his/her income. Hence, such kind of individuals become possible targets to criminals. It was suggested in Meier e Miethe (1993) that the number of components of a family can help diminish the victimization risks. Considering this affirmation, the estimations confirmed this empirical evidence, showing a positive relation impact on repeat victimization.

Gini index has a positive relation with repeat victimization and this effect is the highest among all other variables. However, important sources of variations were not employed in the present study. An important omission is discussion about the separate effects of unobserved heterogeneity and event dependence (see Tseloni e Pease 2004). Such distinction would be crucial for offering a set of more specific recommendation policies.

In face of all that, we can conclude that repeat victimization occurs mainly due to socioeconomic conditions of the individual and the region. Moreover, the possibility of being victimized is bigger for those who had already been victims in the past. Due to this, it is important to develop the notion that repeat victimization reduction is related to the prevention and the reduction of crimes. The results of this work had made evident that such programs cannot be based only on momentary questions, such as increasing the police contingent in the streets, improving its armament and/or constructing new penitentiaries, but also policy makers must pay attention to questions that have long lasting effects, through policies that seek economic and social improvements in the region. In fact, in accordance with Brasil (2003), in order to give an end to criminality and victimization, is necessary to offer security for all the population. Due to this, the government must be committed to create polices that fight against the main variable that lead to the increase of violence. Such variables are mainly centered in the economic, social and public security spheres.

Some important improvements in the paper remain to be done, though. First, the estimation using an updated data set should have the highest priority, since the PNAD 1988 data is a rather old empirical evidence, serving only here as an illustrative exercise. Second, we believe that multilevel model should be an important issue, only touched tangentially by controlling for cluster effects. Given the natural clustering of observation inside the household, and the evident behavioral implications for crime prevention, a count model that incorporates a multilevel (in our case, two levels) is of paramount importance. This is especially relevant because multilevel count data is not a novelty in the criminological community, for instances, see Tseloni e Pease (2003) and Tseloni e Pease (2004). Third, as already mentioned, the issue of unobserved heterogeneity *versus* state dependence was far from complete in our paper. This represents a challenge, however, because we do not have longitudinal observations. Different approaches must be worked out. Finally, cross-effects as well as polynomials effects must be included, specially involving the variables AGE, EDUCATION, INCOME and GINI. However, this is not a trivial task as these effects are not as easily obtained as in the case of linear models. Good references about this very important, but almost entirely neglected, issue are Norton et alii (2004) and Ai e Norton (2003).

References

Ai, C. & Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics Letters*, 80:123–129.

Araújo, A. & Fajnzylber, P. (2001). O que causa a criminalidade violenta no Brasil? Uma análise a partir do modelo econômico do crime: 1981 a 1996. Texto de discussão 162, CEDEPLAR/UFMG.

Barslund, M., Rand, J., Tarp, F., & Chiconela, J. (2007). Understanding victimization: The case of Mozambique. *World Development*, 35(7):1237–1258.

Becker, G. S. (1968). Crime and punishment: An economic approach. *The Journal of Political Economy*, 76(2):169–217.

Blalock, H. M. (1984). Contextual effects models: Theoretical and methodological issues. *American Review of Sociology*, 10:353–372.

Bourguignon, F., Sanchez, F., & Nunez, J. (2003). A Structural Model of Crime and Inequality in Colombia. *Journal of the European Economic Association*, 1(2–3):440–449.

Brasil (2003). Projeto Segurança Pública para o Brasil. Technical report, Fundação Djalma Guimarães.

Brazil (1988). PNAD. National Research for Sample of Domiciles.

Brazil (2008). IPEADATA. Available in: `www.ipeadata.gov.br`. Access in Jan 03 2008.

Cameron, A. C. & Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications*. Cambrigde University Press.

Choe, J. (2008). Income inequality and crime in the United States. *Economic Letters*, 101:31–33.

Dickens, W. & Katz, L. F. (1987). *Unemployment and the Structure of Labor Markets*, chapter Interindustry Wage Differences and Industry Characteristics. Oxford.

Durlauf, S. N. (2004). *Handbook of Regional and Urban Economics*, chapter Neighborhood Effects, pages 2174–2242. Elsevier.

Ehrlich, I. (1973). Participation in illegitimate activities: A theoretical and empirical investigation. *Journal of Political Economy*, 81(3):521–65.

Fajnzylber, P. & et al. (1998). Determinants of crime rates in Latin American and the world, an empirical assessment. Technical report, The Word Bank.

Fajnzylber, P., Lederman, D., & Loayza, N. (2002a). Inequality and violent crime. *Journal of Law and Economics*, 45:1–39.

Fajnzylber, P., Lederman, D., & Loayza, N. (2002b). What causes violent crime? *European Economic Review*, 46(7):1323–1357.

Farrell, G. & Pease, K. (2001). *Repeat victimization*. Criminal Justice Press.

Gaviria, A. & Pagés, C. (1999). Patterns of crime victimization in Latin American. Working Paper 408, Inter-American Development Bank.

Gomes, F. & Paz, L. S. (2007). The determinants of criminal victimization in

São Paulo state-Brazil. Technical Report WPE 25-2007, IBMEC.

Kleemans, E. R. (2001). Repeat burglary victimization: Results of empirical research in the Netherlands. In G., F. & Pease, K., editors, *Repeat Victimization*, volume 12. Criminal Justice Press.

Latinobarometer (1999). Data set. `www.lationobarometro.org`.

Laycock, G. (2001). Hypothesis-based research: The repeat victimization story. *Criminal Justice*, 1(1):59–82.

Levitt, S. D. (1999). The changing relationship between income and crime victimization. *Economic Policy Review*, September:87–98.

MacDonald, Z. (2001). Revisiting the dark figure: A microeconometric analysis of the under-reportingof property crime and its implications. *British Journal of Criminology*, 41:127–149.

MacDonald, Z. (2002). Official crime statistics: Their use and interpretation. *The Economic Journal*, 112:F85–F106.

Meier, R. F. & Miethe, T. D. (1993). Understanding theories of criminal victimization. Technical report, University of Chicago.

Miethe, T. D. & McDowall, D. (1993). Contextual effects in models of criminal victimization. *Social Forces*, 71 (3):74–759.

Moulton, B. (1986). Random group effects and the precision of regresion estimates. *Journal of Econometrics*, 32:385–397.

Moulton, B. (1990). An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *The Review of Economics and Statistics*, 72:334–338.

Norton, E. C., Wang, H., & Ai, C. (2004). Computing interaction effects and standard errors in logit and probit models. *The Stata Journal*, 4(2):154–167.

Pease, K. (1998). Repeat victimization: Taking stock. Series Paper 90, London: Home Offica.

Resende, J. P. (2007). Crime social, castigo social: O efeito da desigualdade de renda sobre as taxas de criminalidade nos grandes municípios brasileiros. Master's thesis, CEDEPLAR/UFMG.

Santos, M. J. & Kassouf, A. L. (2008). Existe explicação econômica para o sub-registro de crimes contra a propriedade? *Economia Aplicada*, 12:5–27.

Sparks, R. F. (1981). Multiple victimization: evidence, theory, and future research. *Journal of Criminal Law and Criminology*, 72(2):762–778.

Tella, R. D., Galiani, S., & Schargrodsky, E. (2002). Crime victimization and income distribution. Technical report, Inter-American Development Bank.

Thacher, D. (2004). The rich get richer and the poor get robbed: Inequality in U. S. criminal victimization, 1974-2000. *Journal of Quantitative Criminology*, 20(2):89–116.

Topel, R. H. (1986). Local labor markets. *Journal of Political Economy*, 94:111–143.

Tseloni, A. & Pease, K. (2003). Repeat personal victimization. "Boosts" or "flags"? *British Journal of Criminology*, 43:196–212.

Tseloni, A. & Pease, K. (2004). Repeat personal victimization: random effects, event dependence and unexplained heterogeneity. *British Journal of Criminology*, 44:931–945.

Wen, M., Browining, C. R., & Cagney, K. A. (2003). Poverty, affluence, and income inequality: Neighborhood economic structure and its implication for health. *Social Science and Medicine*, 57:843–860.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Wooldridge, J. M. (2006). Cluster sample methods in applied econometrics: An extended analysis. Technical report, Michigan State University.

Ybarra, L. M. R. & Lohr, S. L. (2002). Estimates of repeat victimization using the nacional crime victimization survey. *Journal of Quantitative Criminology*, 18(1):1–21.