

HETEROGENEITY IN THE FREQUENCY DISTRIBUTION OF CRIME VICTIMIZATION

Tim Hope and Paul A. Norris¹

October, 2012

Revised, for resubmission to the Journal of Quantitative Criminology

DISTRIBUTIONAL HETEROGENEITY OF CRIME VICTIMIZATION

Please address correspondence to:

Professor Tim Hope
School of Humanities, Languages and Social Sciences
University of Salford
Salford, Manchester M5 4WT UK
t.j.hope@salford.ac.uk timothy.j.hope@gmail.com

¹ Respectively: Centre for Social Research, University of Salford, Manchester, England; and School of Social and Political Science, University of Edinburgh, Scotland. The research for this paper was carried out while the authors were, respectively, Senior Visiting Research Fellow and Research Fellow at the Scottish Centre for Crime and Justice Research (SCCJR), University of Edinburgh, whose support and encouragement is gratefully acknowledged, especially that of Susan McVie and Richard Sparks. The UK Data Archive made data available. We also wish to acknowledge valuable discussions with Kauko Aromaa and Frank Morgan and comments from the anonymous reviewers that have helped to improve the paper. This paper is dedicated to the memory of Alan Trickett.

Heterogeneity in the Frequency Distribution of Crime Victimization

Objectives: Tests the idea that the frequency distribution typically observed in cross-sectional crime victimization data sampled from surveys of general populations is a heterogeneously distributed result of the mixing of two latent processes associated, respectively, with each of the tails of the distribution.

Methods: Datasets are assembled from a number of samples taken from the British Crime Survey (BCS) and the Scottish Crime Victimization Survey (SCVS). Latent Class Analysis is used to explore the probable, latent distributions of individual property crime and personal crime victimization matrices that express the frequency and type of victimization that are self-reported by respondents over the survey recall period.

Results: The analysis obtains broadly similar solutions for both types of victimization across the respective datasets. It is demonstrated that a hypothesized mixing process will produce a heterogeneous set of local sub-distributions: a large sub-population that is predominantly not victimized, a very small 'chronic' sub-population that is frequently and consistently victimized across crime-type, and an 'intermediate' sub-population (whose granularity varies with sample size) to whom the bulk of victimization occurs. Additionally, attention is paid to the position of very high frequency victimization within these sub-populations.

Conclusions: The analysis supports the idea that crime victimization may be a function of two propensities: for immunity, and exposure. It demonstrates that zero-inflation is also a defining feature of the distribution that needs to be set alongside the significance that has been attached to the thickness of its right tail. The results suggest a new baseline model for investigating population distributions of crime victimization.

1 Background

1.1 Introduction

Previously in this *Journal*, McDowell (2010) has reminded us of the very common shape of the individual-level frequency distribution observed amongst general populations over a wide range of criminological outcomes. The distribution of property crime victimization taken from the Scottish Crime Victimization Survey, illustrated in Figure 1, is a typical example². A concern to develop and test explanatory models, often using linear regression equations to estimate conditional means of the dependent variables of interest, encourages transformations and other model adjustments that seek to remove the complexities of statistical analysis suggested by this kind of skew. Yet:

“...No approach has seriously considered the source of the skews...or made their properties objects of study in themselves. The existing methods mould the data to meet model requirements, and researchers choose them as a matter of convenience. Yet asymmetry is ubiquitous in criminology and avoiding it amounts to disregarding an obvious fact”. (McDowell, 2010: 433).

This paper seeks to remedy this situation a little by exploring the probable, unobserved latent distribution of crime victimization; that is, to investigate the statistical *data generating process* (DGP) that might be producing the observed distribution.

Figure 1 around here

1.2 The problem of over-dispersion

Typically, the frequency distribution of crime victimization is over-dispersed; that is, the sample variance exceeds its mean. There are two possible reasons for over-dispersion (Cameron and Trivedi, 1998): on the one hand, it might arise because of the ‘excessive’ proportion of zeros to be found at the ‘left-hand tail’ of the distribution (in this case showing that the majority of the sampled population do not report incidents over the recall period). On the other hand, over-dispersion in the distribution could be caused by the ‘thickness’ of its ‘right-hand tail’, produced by an excessive prevalence of cases with higher frequencies of self-reported victimization incidents (Ellingworth *et al.*, 1995). As Figure 1 illustrates, both seem to be present in the distribution of crime victimization but it is not known which of the two exerts the most influence on its dispersion. It is unfortunate that this question has been overlooked because a better understanding of the underlying probability structure that gives rise to characteristically over-dispersed crime victimization frequency

² Definitions of the crime categories used in this paper are given in the Appendix.

distributions might lead to further progress in understanding the causes and remedies of crime victimization. Thus the aim of this paper is to explore the latent processes that might be giving rise to over-dispersion in the frequency distribution of crime victimization.

The standard statistical model for count data, the *Simple Poisson* (SP), is not appropriate for the distribution of crime victimization since its premises are violated by over-dispersion³. Early research suggested that there was a need for an alternative model specification that would account for over-dispersion by allowing for the possibility of there being a heterogeneous mixture of sub-population frequency distributions that together comprise the overall distribution (Nelson, 1980; Sparks, 1981; Sparks et al., 1977). The standard parametric model used to account for over-dispersion that might have resulted from mixing is the *Negative Binomial* (NB), which early research found to have "...an astoundingly good fit" to the distribution of crime victimization in sample survey data (Nelson, 1980: 872). The *Negative Binomial-2 Model* (NB2) gives a better fit to the data than the SP specification (Osborn and Tseloni, 1998), and both multivariate and multi-level models with NB2 specifications have been fitted to British Crime Survey (BCS) and National Crime Victimization Survey (NCVS) data (Tseloni; 2006; Tseloni and Pease, 2004; Tseloni and Pease, 2003; Osborn and Tseloni, 1998)⁴.

Although the NB2 model gives estimates not only of mean effects but also of the extent both of explained and unexplained heterogeneity amongst individual cases, its capacity to account for over-dispersion is limited because it cannot provide information on the particular mixture of sub-distributions within the overall distribution. Although the NB2 regression model specification achieves a better estimate of the effect of population heterogeneity in cross-sectional data, neither can it estimate the likelihood of stochastic components being present in the mixture, such as the existence of a tendency towards repeated victimization (Tseloni and Pease, 2003; Osborn and Tseloni, 1998)⁵. In sum, although a consensus seems to have been reached that the Negative Binomial distribution provides the best fit to

³ The SP model is a benchmark for counts data that consist of a number of discrete events occurring at the unit-level over a fixed time interval (Cameron and Trivedi, 1998). It assumes that successive events for any individual case occur independently of each other over time at a constant rate, an assumption violated by over-dispersion (Nelson, 1980: 871).

⁴ NB2 assumes a Poisson distribution, with gamma-distributed unobserved individual heterogeneity (Cameron and Trivedi, 1998: 71).

⁵ Even the relatively few analyses of longitudinal crime victimization data to date have been unable to distinguish clearly and conclusively the mixing components of the DGP, or the relative importance of heterogeneity versus state-dependency (Bottoms and Costello; 2009; Hope and Trickett, 2008; Hope and Trickett, 2004); Wittebrood and Nieuwbeerta, 2000; Lauritsen and Davis Quinet, 1995). Morgan (2007) provides a detailed discussion and analysis of time-dependent models of crime victimization.

the population frequency distribution of crime victimization, interpreting the results of NB models has proven difficult and uncertain. In large part, this is because the standard model can support several equally plausible interpretations of the nature of the mixture and the resulting heterogeneity of the distribution⁶. In particular, the NB2 model fitted to crime victimization data cannot address a question raised early on in the empirical analysis of crime victimization survey data; namely, to ascertain the relative importance to be attached to the influence of one or the other tails of the distribution (Sparks et al. 1977).

1.3 Researching the distribution of crime victimization: a tale of two tails

Empirical research to date has tended to place its weight on either one or the other tail of the distribution. Initially, attention was given to understanding the prevalence of victims within the population, with an emphasis on finding those factors that distinguish the part of the population that has experienced crime victimization from the other part that has not (Nelson, 1980). First exploited by Hindelang, *et al.* (1978), the approach sidesteps the need to account for over-dispersion by ignoring it (Sparks, 1981). Instead, the observed frequency distribution is truncated into two discrete outcomes, non-victims (indicated by zero values), and a conflation of all other positive count values into the category of 'victim'. The probability of membership in the victim category relative to the non-victim category is then estimated by model covariates (Osborn et al., 1992). The weight of evidence shows that heterogeneity in the probability of crime victimization is likely to be attributable causally to variables directly or indirectly measuring individuals' life styles, routine activities and residential circumstances, all of which are held to serve as indicators of the likelihood of exposure to potential offenders (Tseloni, 2006; Miethe and Meier, 1994).

An approach that focuses on prevalence tends to be taxonomic, regarding the cut-points between categories (usually between the zeros and the positive counts) as the essential, defining characteristic of the distribution, and the source of its heterogeneity. Osborn *et al.* (1996) extend this approach by fitting a conditional *double-hurdle* model for household property crime victimization, comparing the transition from non-victim to victim household (the first hurdle) with the subsequent transition to a higher-frequency victimization likelihood (the second hurdle)⁷. The results demonstrate the explanatory insignificance of the second hurdle (indexing the frequency component) and its dependence upon the first hurdle (indexing the prevalence component). Nevertheless, although these results support an interpretation of the distribution as defined by a bifurcation between two distinct groups – that is, 'all non-victims' *versus* 'all-victims' - it still remains unclear whether

⁶ The NB model specification can support a variety of probability mechanisms, a range of mixtures that might produce it (Cameron and Trivedi, 1998: 102) and a variety of Mixed-Poisson models to help explain it (Karlis and Xekalaki, 2005).

⁷ Specifically, the study utilized a bivariate probit regression model with censoring (Osborn *et al.*, 1996).

it is prevalence (shaped by the zeros) or frequency (shaped by the positive values) that exerts the defining influence.

A general tendency towards non-victimization (producing an excess of zeros) as a reason for over-dispersion has been examined in a panel study using probabilities of subsequent victimization conditional upon varying levels of prior victimization (Hope and Trickett, 2008)⁸. Support was found for four related processes:

- i. A general tendency toward non-victimization over time, regardless of prior level of victimization.
- ii. A tendency for Non-victims to retain their (non-victimization) state over time;
- iii. A tendency also for High-level victims to retain their high-frequency state over time.
- iv. A tendency for Low-level victims not to retain their state, reverting to the general tendency towards non-victimization.

A general tendency towards non-victimization over time has also been detected in police-recorded crime data sampled in the English city of Sheffield (Bottoms and Costello, 2009) and in Perth, Western Australia (Morgan, 2007). In contrast, then, to an explanation of the data generating process of crime victimization as characterized solely by exposure, a tendency towards non-victimization over time also suggests a complementary process of 'non-exposure' or *immunity*, albeit conditioned by prior state (Hope and Trickett, 2008). Again, the distribution appears to be heterogeneous, although this time from the perspective of immunity; not only does the state of non-victimization appear to be relatively stable but so too does the state of high-level victimization, albeit as a significant deviation from the general tendency towards non-victimization.

An alternative approach has sought to investigate the process that accounts for a distribution of victims with varying levels of incidents amongst whom the sum of crime victimization is distributed disproportionately. A primary motivation is to understand the contribution of multiple victims to the overall crime (incidence) rate (Pease and Farrell, 2007; Farrell and Pease, 2007; Tseloni and Pease, 2004; Tseloni and Pease, 2003; Tseloni et al., 2002; Hope, et al., 2001; Farrell and Pease, 1993). Proponents of this view tend to rely upon the practical justifications of crime prevention (Pease, 1998; Farrell, 1995), arguing for attention to be paid to the 'power few' (Sherman, 2007), that is, the infrequent number of victims who appear to exert disproportionate power over the victimization frequency distribution. So-called 'power laws' are cited, such as the 'Pareto 80/20 Rule', as a reason for

⁸ The categories of counts were: Non-victim (no incidents over the recall period); Low-level (1 victimization incident over the recall period); and High-level (2 or more incidents). The waves were measured by a panel embedded in two cross-sectional self-report sample surveys, with 12-month recall periods, separated by a period of three years between each sampling (Hope and Trickett, 2008: 47; see also Hope, 1995).

concentrating upon the minority (i.e. approximately 20 per cent) who appear to be generating the majority of the count at issue (i.e. approximately 80 per cent) (Eck et al., 2007). Explicitly, Tseloni et al. (2010) cast this as a resource allocation problem for the targeted deployment of policing and other crime-preventive resources.

The focus on frequent victims has assumed the existence of a stochastic process of repeated crime victimization, particularly, the possible contagion of crime victimization itself (Hope, 2007a); that is, a view that the probability of present victimization can be best predicted from the previous victimization of targeted individuals and households or their proximate neighbors (Pitcher and Johnson, 2011; Short et al., 2009). The stochastic element might be called more appropriately 'event-dependency', since contagion might be transmitted via incident-specific mechanisms – i.e. properties attributable to the incident, including the inferred crime-selection activities of offenders (Pease, 1998: 8). Nevertheless, not only is prior victimization associated with contemporaneous victimization, suggesting repeat victimization within crime types (Ellingworth et al., 1997), but there are also prior effects on contemporaneous victimization across crime-types, as well as associations between contemporaneous crime types; evidence suggesting the possible existence of a discrete category of multiple victims (Hope *et al.*, 2001)⁹.

Thus, support for the idea that the distribution of crime victimization counts is heterogeneous emerges whether the perspective focuses on the question of victimization-frequency (Tseloni, 2006) or on that of victim-prevalence (Osborn *et al.*, 1996). As such, although both would seem necessary perspectives in order to grasp the nature of the heterogeneity of the data generating process, it remains undecided how much emphasis to place either upon one or the other. This problem is analogous to understanding the epidemiology of a disease: on the one hand, should we be concerned with differences amongst the hosts (i.e. victims) in their susceptibility to the disease (i.e. crime victimization), or should we be concerned with the communicable contagiousness of the disease itself (i.e. from one incident to another)? Since both seem important, it might be useful to counterpoise victim-prevalence and victimization-frequency as the two necessary ingredients of the distribution of crime victimization. That is, to characterize the latent distribution of crime victimization as comprising the twin processes of *immunity* (associated with victim-prevalence) and *exposure* (associated with victimization-frequency), which mix together to generate a population distribution of crime victimization that is characteristically over-dispersed.

⁹ In addition to having been victimized over the past five years, contemporaneous multiple crime-type victims were significantly more likely to be younger adults, living with children, renting from the social or public housing authorities and living in poorer, urban areas; all characteristic of economically marginal and socially vulnerable sub-populations in the UK who might be more likely to be most often exposed to active offenders (Hope *et al.*, 2001).

1.3 Conceptualizing the mixture distribution of crime victimization.

The distributions in question consist of counts of crime victimization events experienced by individuals over periods of time. Given that research has found evidence of multiple crime victimization both within and between different types of crime, we might conceive of each 'target individual' as having their own specific *crime victimization event matrix*, comprising counts in three dimensions: time (e.g. a recall period in a survey context), frequency (the number of counts over that time-period), and type (the variety of different types of event) (Hope et al., 2001). This paper is concerned to describe the characteristics of the distribution of these individual event matrices amongst the population¹⁰.

At any moment within the time-period covered by the event matrix, a 'target' of crime can experience only one of two outcomes - no victimization (0) or victimization (1). Thus the true distribution of target population event matrices might be the product of a mixture of two probability distributions: the first being the probability of each case being a zero (which could be taken to correspond to the process of immunity), the second being the probability of a positive (non-zero) count (corresponding to the process of exposure). As suggested by Figure 2, each target individual has two propensities (i.e. positions on the respective probability distributions), one for a victimization event to occur, and one for it not to occur¹¹. Empirically, these propensities may combine (or mix) together into a single probability distribution. Like the cat in Schrödinger's thought experiment in quantum theory, a victimization event at the point of occurrence (i.e. mixing) has the propensity either to be or not to be. Indeed, that is the question of explaining the observed distribution of crime victimization: given these two propensities, how do the two propensities for immunity and exposure mix together to produce a distribution of crime victimization outcomes that has the statistical properties of over-dispersion?

Figure 2 about here

¹⁰ Tseloni *et al.* (2010) also call attention to *composite crimes*, i.e. when more than one crime type coincides as part of a single event. This might be a further dimension to an individual's crime victimization event matrix but is excluded from consideration here both on the grounds of a lack of accessible data on event-composition (Tseloni *et al.*, 2010) and because its inclusion would increase the complexity of subsequent analysis (probably exponentially) to the detriment of exposition. Still, we should concede that all efforts to classify crime victimization events or states statistically are highly circumscribed in their phenomenology, due largely to the constraints of the incident-counting machinery, whether sample surveying or offence recording (Biderman, 1981).

¹¹ These two hypothetical probability distributions may be thought of as propensities, in the sense of Popper's *propensity interpretation of probability* (Popper, 1983/1957); that is, as heuristics that give an idealized account of the objective, relational properties of the physical world (Gillies, 2000).

The composite frequency distribution of crime victimization events, defined as individual crime victimization event matrices aggregated to a portmanteaux variable (y_i), can be conceptualized as a mixture model, whose distribution is specified as $(y_i | \mathbf{x}_i, v_i)$, where y_i is conditional both on observable covariates \mathbf{x}_i and on an unobserved heterogeneity term v_i for observation i , with the marginal distribution obtained by averaging with respect to v_i . This has a functional form of an exponential mean with multiplicative error; that is, $E[y_i | \mathbf{x}_i, v_i] = \exp(\mathbf{x}_i' \beta) v_i$. The concern of this paper lies in the latent heterogeneity term (v_i). This is considered as a multiplicative product, representing the interaction between the two components of immunity and exposure, where the stochastic term (v_i) is independent of the regressors (Cameron and Trivedi, 1998).

The assumption that the unobserved heterogeneity of the distribution (v_i) is multiplicative is important in accounting for the specific statistical properties of over-dispersion. Heterogeneity is not necessarily implied by over-dispersion because processes other than heterogeneity, for example state-dependency, can also produce over-dispersion, though without heterogeneity. For instance, the distribution of crime victimization might merely be a product of the exposure process alone, discounting the role of excess zeros, as is implied in the interpretation of the distribution as a product of event-dependency¹². Nevertheless, the presence of excess zeros *necessarily* implies heterogeneity (Mullahy, 1997)¹³. Mixing based specifically on multiplicative heterogeneity will produce both of the phenomena of over-dispersion: not only does mixing cause the proportion of zero counts to increase, but the mixture also exhibits a thicker right-tail, relative to the Simple Poisson (Cameron and Trivedi, 1998: 99)¹⁴. Empirical confirmation of this result for the distribution of property crime victimization in BCS data comes from Osborn and

¹² For example, Tseloni et al. test a *Bivariate Zero-inflated Poisson* (BZIP) model that assumes zero counts to arise in two ways: first, that “...no crime occurs, with probability p and produces only zeros, while the other state where crime exists, occurs with probability $1 - p$ and leads to a standard Poisson count” (2010: 334).

¹³ A substantial, significant zero-inflation coefficient is identified in the BZIP model estimated by Tseloni et al (2010).

¹⁴ The idea that distributions (such as crime victimization) are multiplicative heterogeneous mixtures is supported by reference to the Two-Crossings Theorem (Mullahy, 1997): “...*Two Crossings Theorem*. For the random variable y , continuous or discrete, let $f(y | \mathbf{x}, v)$ denote an exponential family conditional (on v) model density and let $E[v] = 1$, $V[v] = \sigma^2 > 0$. Then the mixed (marginal with respect to v) distribution $h(y | \mathbf{x}) = E_v f(y | \mathbf{x}, v)$ will have heavier tails than $f(y | \mathbf{x}, v)$ in the sense that the sign pattern of marginal minus the conditional $h(y | \mathbf{x}) - f(y | \mathbf{x}, v)$ is $\{+, -, +\}$ as y increases on its support. That is, for the same mean, any marginal distribution must ‘cross’ the conditional distribution twice, first from above and then from below, the first crossing accounting for a relative excess of zeros, and the second for the thickness of the right tail” (Cameron and Trevedi, 1999: 99).

Tseloni (1998) where both properties were observed in the fitted NB2 distribution compared to an SP regression model with the same covariates.

The aim of this paper is to infer the latent heterogeneity (v_i) of the data generating process that produces the observed distribution of crime victimization amongst the target population as a mixing of the processes of immunity and exposure. We characterize the outcome distribution (y) as a *finite mixture* of C sub-populations: where (y) is thought of as a draw from a super-population that is an additive mixture of C in proportions π_1, \dots, π_C , where

$$\sum_{j=1}^C \pi_j = 1, \pi_j \geq 0 \quad (j = 1, \dots, C) \quad [1]$$

(Cameron and Trivedi, 1998: 128).

In the analysis that follows, we use the technique of *Latent Class Analysis* (LCA) to estimate $\sum_{j=1}^C \pi_j$.

Contemporary usage of LCA in criminology tends to be taxonomic; for example, to identify sub-population groups with different criminal career trajectories (Brame et al., 2006). Nevertheless, confusion (and controversy) has arisen regarding the interpretation of the results from LCA in this context: on the one hand, there is the criticism that group-based modeling bestows upon the groups that are identified a spurious substantive reality since they may just be artifacts of the method used to produce them (Sampson and Laub, 2005; 2003); on the other hand, there is the view that the particular groups identified by LCA are in themselves of less substantive importance than is the depiction of the distribution that LCA reveals (Nagin and Tremblay, 2005). In the present context, a similar confusion might arise; namely, whether the true distribution of crime victimization would be best characterized as a single, continuous and possibly stochastic dimension of crime victimization risk, or whether it is a composite of discrete sub-populations (types) of crime victims.

By itself, the technique of LCA cannot resolve this issue because the problem is at root epistemological rather than methodological; the particular interpretation adopted depending upon the use to which the LCA is put (Pickles and Angold, 2003). Nevertheless, even if it does not offer decidability, finite mixture methodology has practical advantages, especially for exploring the nature of heterogeneity given the problem at hand: on the one hand, LCA produces a representation of a smaller number of latent classes that could be considered as natural groupings, if that is what prior theory expects. For example, there might be a discrete group of multiple victims, or that the distinction between victims and non-victims is decisive. On the other hand, a finite mixture may be a useful way of "...flexibly and parsimoniously modeling the data, with each mixture component providing a local approximation in some part of the true distribution" (Cameron and Trivedi, 1999: 129).

Our logic is probabilistic (Hacking 2001) and confirmatory (Howson and Urbach, 1989). We employ LCA to estimate a parsimonious set of *ordered latent classes* that are expected to (1) replicate the observed distribution, particularly its over-

dispersion, while (2) revealing the latent processes of immunity and exposure. Logically, the approach is similar to a *Bayesian Network* approach (Pearl, 2000:7.1.4): two identical models of the outcome distribution are postulated, with each data generating possibility acting as a counterfactual to the other; the one representing the distribution as if it were determined solely by victim-prevalence (immunity), the other as if it were determined solely by victimization-frequency (exposure), as schematized in Figure 2. This conception allows a latent distribution to be estimated by local computational methods (Pearl, 2000: 214), dovetailing nicely with the capacity of finite mixture models to approximate latent continuous distributions¹⁵.

2. Data and Method

2.1 Data sources

Specific datasets were created for the analysis compiled from several waves of the British Crime Survey (BCS) and the Scottish Crime Victimization Survey (SCVS)¹⁶. Respondents from England and Wales (taken from the BCS) and Scotland (taken from the SCVS) were treated separately. We take advantage of the fact that these different but closely related sample surveys cover separate jurisdictions in order to provide some grounds for generalizing our findings. Property crime victimization also needs to be modeled separately from personal crime. In both the BCS and SCVS, respondents report on the property crime experiences of their household, while reporting on their own experiences with regard to personal crime. Consequently, different weighting variables, reflecting both this distinction and other sample design features, need to be used for analyzing property and personal crime respectively. Further, precautions needed to be taken to cope with the effect of certain aspects of both sample selection and response error inherent in the data.

2.2 Correcting for sample selection bias

Both the BCS and SCVS have been repeated several times over the past twenty-five years (both surveys are now conducted continuously). Only for the first BCS in 1982 was data collected about victimization in (southern) Scotland as well as for England and Wales. Since then, separate series of BCS and SCVS surveys have been conducted (with the SCVS now covering the whole of Scotland). Both surveys have used clustered multi-stage sample designs, which have varied over time. Recognizing

¹⁵ Pickles and Angold (2003) describe the approach as akin to *Gaussian Quadrature*, that represents "...smoothly varying densities of a distribution by estimating a limited number of spikes at a set of specific values...each spike is assigned a probability weight...the spikes are, mathematically speaking, identical to a set of ordered latent classes" (Pickles and Angold, 2003: 540).

¹⁶ recently, both Survey series have undergone name changes, while the SCVS has had several changes of nomenclature. These names are retained however as those most appropriate at the time the surveys were conducted.

this, recommended sample weights (used to correct for bias against population representativeness due to the complex survey design) were applied to the data input to our dataset (see e.g. Bolling, 2007; Brown, 2007).

Combining responses from several surveys conducted since the early 1990s created the datasets used in this paper.¹⁷ Combining several years' worth of data increases the sample size available for analysis. This is potentially useful because victimization is generally a rare event (indicated by the 'excess zeros') and a large sample size may therefore be necessary to capture the wide range of different individual event matrices. Combining responses from different surveys also could be expected to reduce the risk of any bias, or context-specific results, which relate to the data collected in any given year (for instance due to the use of a particular sample frame, or change in sampling strategy). For instance, the sample size of the BCS has generally increased over time, and underwent a major re-design, effective from 2000 so that the latter two sweeps in our source data (2003/4 and 2006/7) differ from those earlier in the series¹⁸. In view of this, new weight variables proportionate to overall sample size were calculated so that each year's data contributed 20% of the overall sample when weights were applied. This allowed the full sample from each year to be analyzed but ensures that the results are not unduly influenced by data from any one particular year¹⁹.

The necessary use of weighting to correct for sampling bias raises a basic dilemma with regard to the secondary analysis of data drawn from complex sample designs, since many parametric tests cannot be applied to weighted data. In particular, 'bootstrap' likelihood ratio tests cannot be used in LCA with weighted data, even though they may perform better in determining the most appropriate number of classes, particularly producing solutions that are less dependent on sample size²⁰.

¹⁷ Responses for England and Wales were taken from the 1992, 1996, 2001, 2003/4 and 2006/7 BCS. Responses for Scotland were taken from the 1993, 1996, 2000, 2003 and 2006 SCVS.

¹⁸ Chiefly, this has been done to facilitate the development of performance indicators to assist the Home Office in its governance of the police service in England and Wales (Allen, 2007).

¹⁹ Further comfort that the different sample designs of the components of our dataset have not biased our results is provided by a recent and extensive investigation of this issue, concluding that, under all of its sample designs, the BCS has generated estimates of victimization with low levels of sampling error (Tipping et al., 2010).

²⁰ Methods based on simulation have been suggested as providing a more accurate indication as to how many classes are needed for an LCA model to accurately reflect an underlying dataset (Nylund et al, 2007). Of these, the most common is the Bootstrap Log-likelihood Ratio Test (BLR) (McLachlan and Peel, 2000). This calculates a p-value that indicates whether or not the inclusion of an additional class significantly improves the relationship between the data and the model. However,

Nevertheless, mindful of McDowell's (2010) criticism voiced at the start of this paper, we are anxious not to overly 'bend' our analysis or massage our data to suit our preferred technique. Further, the main purpose of our analysis is to use LCA not for the purpose of estimating the prevalence of latent groups in the population but as a method for investigating a latent and perhaps continuous distribution (Pickles and Angold, 2003), where precision of the estimates of the 'spikes' is arguably of lesser importance than the estimation of the relationship between them.²¹ Finally, there is the problem of interpretation: how would one interpret the results that were obtained by bootstrap methods applied to unweighted data, compared to standard methods applied to weighted data: if they differed, would this reflect the weighting or the method; but if they did not, why then sacrifice representativeness for the sake of precision? Consequently, we use standard rather than bootstrap methods, relying on the large size of our dataset.

2.3 Correcting for response error

There are two sources of data in the BCS and SCVS from which a count of respondents' experiences of crime victimization might be derived. First, there are the 'screener questions' that appear near the beginning of each crime survey questionnaire. These are the victims' self-reported estimate in response to a query about the number of times various incidents (such as those listed in the Appendix) have been experienced over the recall period. The screener questions consist of a simple count. In the official crime counting procedure for both surveys, the screener questions provide merely the baseline for the construction of the counts that make up the estimates of prevalence and incidence rates. The mechanism for producing the latter is through the completion of 'victim forms', which comprise the second source of data on crime counts. These ask a number of descriptive follow-up questions about incidents that assist in the calculation of the overall prevalence and incidence rates as well as providing detailed information on the circumstances of each event. Each victim form is counted as a single incident unless the respondent answers a question as to whether a similar sequence of incidents had occurred over the recall period, probably by the same perpetrators; if so, the number of these is recorded on the relevant victim form, and these are then counted as the number of incidents that comprise a 'series' of crime victimization events²².

the use of these methods with weighted data is not well understood or widely implemented at this time.

²¹ See Fn 14. Morgan (2007) makes an analogous point regarding the merits of event-hazard models for estimating stochastic crime victimization processes.

²² Although the counts from series that are derived from victim forms are subsequently capped at 6, being the maximum count per series per victim form (Farrell and Pease, 2007).

This crime counting procedure differs somewhat from that used in the NCVS (Planty and Strom, 2007); in the BCS and the SCVS, the victim forms comprise a selection taken from the positive response to the screener questions, with the addition to the count that is derived from series of crimes taken from a question asked only of those respondents who complete victim forms. Whereas the count taken from the screener questions counts the whole sample, the count taken from the victim form (including the estimate of series) is censored by the selection procedure for completing victim forms. Furthermore, the use of the count derived from the victim form may entail a selection bias vis-à-vis the count obtained from the screener questions. Not only are respondents restricted in the number of victim forms they are asked to complete (up to a maximum of six per victim) but also these are a non-random selection from the number reported in the screener questions, since the more serious and less common types of victimization are followed-up first.

Although the victim forms yield a specific count of series of incidents while the screener questions do not, in view of how the victim forms are selected it is less clear whether a count based on victim forms would come out the same as a count derived from the screener questions. Nevertheless, since the screener questions are asked of the whole sample, prior to the administration of the victim forms, and also ask for an unqualified response, they ought not to provide a lesser count than that derived from the victim form. In contrast, the victim form count might give a lesser count than the screener questions in view of the selection procedure of the victim form, which could result in some events that are counted in the screener questions being disregarded. Even so, just as it is unknown what proportion of counts from series are counted in the screener questions, so it is also unknown what proportion of counts from the screener questions would be defined as emanating from series. Still, it may be safer to use the screener questions to generate the counts to be used in this analysis rather than risk the possible selection bias of the victim forms, especially as the concern is to model respondents' pattern of counts of their crime victimization experiences. For these reasons, previous research has preferred to use the screener questions when modeling individual-level victimization²³ as well as for describing the frequency distribution²⁴.

Some concern has been expressed that counting practices that ignore the incidence of series greatly underestimate not only the count of crime victimization (Planty and Strom, 2007) but also the contribution that so-defined repeat victims might make to it (Farrell and Pease, 2007). Be that as it may, this issue is not strictly relevant to the

²³ For example, Tseloni, 2006; Hope et al., 2001; Osborn and Tseloni, 1998; Ellingworth et al., 1997; Osborn et al., 1996; Osborn et al., 1992; Trickett et al., 1992.

²⁴ "...repeated victimizations cannot be looked at comprehensively through the victim forms, because of the constraint imposed on the number of victim forms completed per respondent, and the maximum number of events 'permissible' in the series...there is no alternative to the use of the main questionnaire for the purposes advanced here...The general pattern is robust across different limits to the number of victimizations allowable per person" (Ellingworth, et al., 1995: 361).

concern of this paper, which is rather to analyze the distribution of individual respondents' crime victimization event matrices, which might be estimated more reliably using the screener questions. As such, the risk of underestimating the contribution to crime victimization counts of 'series' of incidents in the subsequent analysis becomes irrelevant, since the definition of a series is a function of another data collection procedure in these surveys (i.e. the victim form) that is not a property of the procedure (i.e. the screener questions) from which data to be used in this analysis are derived.

The BCS and SCVS each contain six screener questions that concern property crime and four questions that are concerned with personal victimization. These are listed in the Appendix to this paper. Question wording is very similar and we have used the SCVS questions for illustration. In addition, both surveys include questions about sexual violence and violence in the home. However, these latter questions were not always asked of all respondents (for instance in some years the sexual violence questions were only asked of female respondents), and were therefore excluded on the grounds of consistency in the analysis.

The typical distribution of crime victimization (as in Figure 1) suggests that there will be a small proportion of victims who experience a high number of incidents. However, problems of statistical estimation arise if the ratio between the two is large (another problem posed by over-dispersion). With LCA modeling this can prevent model convergence, or may lead to models where the extreme cases are separated from the rest of the data, at the expense of under-estimating heterogeneity in the remainder of sample; in other words, where the division between the few extreme values and the rest of the population itself becomes the defining feature of the distribution. Although this latter possibility may be of some interest (Farrell and Pease, 2007), in practice it conflicts with the aim of this paper, which is to investigate heterogeneity occurring across the whole distribution.

The two data-columns made available for the screener questions in the questionnaire would allow respondents potentially to report in excess of 90 incidents of each type of victimization. It was therefore necessary to cap the values at a reasonable maximum to limit the possible problems that might otherwise arise in the analysis. Accordingly, the maximum count for each type of victimization was capped at six incidents per crime per respondent. For the twenty crime types considered from the two surveys, the maximum proportion of capped cases was 0.47 per cent (for receiving threats of violence in the BCS), while for 11 crime types the proportion of capped cases was 0.01 per cent or less²⁵.

A further reason for capping individual frequencies is that the data suggest that some who are frequently victimized may have had difficulty in calibrating the frequency precisely, stemming from the sources of error associated with the recall

²⁵ A table summarising the operation of capping is available from the authors, on request.

methodology used in crime victimization surveys (Cantor and Lynch, 2007). For instance, for the personal crime type 'Received threat of violence', which had the most number of respondents reporting more than 6 incidents in the BCS data set (about 0.47% of the victims of this crime type), just under 15% claim to have experienced '10' incidents, 12% '12' incidents, and around 20% reporting either '20', '30', '40', '50' or '60' incidents (but not other multiples). This suggests that the figures provided by respondents having very frequent victimization might be broad estimates at best, using shorthand heuristics to aid their recall, such as ordinal rankings or everyday ways of talking about frequency, for instance, 'once a month'. These may then have been translated into integers during data coding. In addition, there are also victims who respond 'too many to recall'²⁶. Rather than treat these as missing data, reducing estimates of high frequency victims, they have been retained and recoded to the maximum value (of 6). In sum, although the capping of offence frequencies at a maximum of six means the analysis may under-estimate high frequency victimization, it would be more unreliable to do anything else. Nevertheless, we also provide some information on the consequences of capping in our subsequent analysis.

Table 1 gives descriptive statistics for each of the datasets employed in this paper. All the variables considered (irrespective of crime type or jurisdiction) in this analysis exhibit similar characteristics, being highly skewed with in-excess-of 95 per cent of respondents reporting no victimization (for any given crime type), and variances which are substantially larger than their means. Descriptively, these statistics suggest that the distribution of victimization generally follows a zero-inflated negative binomial distribution and it is this that is used as the basis for the subsequent LCA²⁷.

Table 1 about here

2.4 Method

²⁶ Within the BCS and SCVS datasets, such responses are coded with a very high value, typically '97', though this is usually understood as a missing value code (Bolling et al 2007: 71; Brown, 2007: 91). Although these can form a substantial proportion of the cases with values over six (the highest, at 52%, being for the BCS property crime 'entered property and committed theft'), the proportions reporting 'too many to count' vary not only between crime types but also between the BCS and SCVS samples, suggesting a complex outcome of a combination of reporting and coding processes. A tabulation of these cases is available from the authors on request (see also Fn. 23).

²⁷ Models were also created based on negative binomial distributions without zero inflation and, in the case of the BCS data, using a censored distribution capped at 6. While the exact model fit statistics for these different distributions varied, the number of groups identified as optimal for capturing patterns of victimization was the same as when using zero-inflated distributions.

In estimating an LCA model, the class to which an individual belongs is treated initially as missing data (which, in this case, can be estimated based on their reported experience of victimization)²⁸. As LCA is designed to work with incomplete information, an advantage of this technique is that individuals who have answered some but not all of the relevant survey questions can still be included in the model. An LCA model estimates two types of parameters. The first of these estimates the probability of each individual appearing in each class identified by the model and provides a guide as to which class within the model best represents an individual's overall experience of crime. The second set of estimates provides an overview of the average level and mix of victimization experienced by respondents within a particular class. A further advantage of LCA models compared to more traditional clustering methods is that they can handle observed variables that follow non-normal distributions. In view of the discussion presented in the previous section, all the observed variables considered in this analysis were treated as following the zero-inflated negative binomial distribution (see also Table 2, below).

When estimating LCA models, one concern is that the characteristics of the final solution are dependent on the starting values used to initialize the procedure (Muthen, 2001). Finding a latent class solution that can be reached from a range of different starting values (known as global maxima) is therefore seen as an indication that the solution is robust and not simply a function of the starting values employed. Each LCA model in this paper is based on 2,500 sets of starting values²⁹.

There is a long-running debate as to how to identify the optimal number of classes needed to represent the underlying patterns accurately within a dataset. Within Criminology (and applied social statistics more generally) the *Bayesian Information Criterion* (BIC) (Schwarz, 1978) is a commonly used indicator for deciding on how many classes to include in an LCA model (Brame et al., 2006). Based on a model's log-likelihood value, the BIC penalizes models which contain more classes, meaning it may increase as more classes are added; the model providing the lowest BIC statistic is considered the optimal solution, best summarizing differences within the data. Although the BIC has been shown to perform well across a variety of models, incorporating different sample sizes and variables which follow different underlying distributions (Nylund et al, 2007 p557), it may over-estimate the number of groups present in very large samples. An alternative to the BIC is the *Adjusted Bayesian Information Criteria* (ABIC), which incorporates an alternative metric for sample size, and is intended to favor more parsimonious solutions (i.e. those with fewer classes) when using a large sample size. The available evidence suggests that the ABIC performs well across a wide range of datasets, although it may still over-estimate the optimal number of classes when using particularly large samples (Yang, 2006 and

²⁸ Extensive introductions to LCA can be found in Magidson and Vermunt (2004), and McCutcheon (1987).

²⁹ The LCA models in this analysis were estimated using MPlus Version 5.1 (Muthen and Muthen, 1998-2008).

Heterogeneity in the Frequency Distribution of Crime Victimization

Nylund et al, 2007). Reflecting these findings, it is the ABIC that is reported in this paper, especially since the objective is to use classes to estimate the shape of the whole distribution rather than to identify the point estimates of particular classes.

3. Results

3.1. Estimating the distribution from a set of ordered latent classes

The first set of analyses fits a parsimonious set of ordered latent classes (Sections 3.1 – 3.3). It describes the characteristics of the classes in order to test the basic proposition that the observed frequency distribution of crime victimization (y) is generated from a latent mixture (v) that might result from the latent data generating processes of exposure and immunity. The iterative process of identifying latent classes (measured by changes in the ABIC values between solutions) is illustrated in Figures 3 and 4 for property and personal crime, respectively, comparing the BCS with the SCVS. First, it is evident that latent classes can be identified in the distribution of crime victimization, thus suggesting the possibility of heterogeneity in the distribution. Second, this would seem to be a general result, since the iteration process is very similar across all samples³⁰. Third, by far the greatest reduction in the ABIC is obtained by the initial formation of two latent classes. In a context such as this, where measurement is by positive integers, a two-class solution replicates the cut-point between all positive values and zero. As such, the magnitude of the reduction in ABIC for the two class solution, relative to the reductions obtained in subsequent iterations, confirms the importance of the victim-prevalence component of the distribution.

Figure 3 about here

Figure 4 about here

Descriptive data on class membership are presented in Figures 5 – 8. The classes are ordered along the x-axis in terms of the proportion of the sample they represent. Regardless of crime type or sample, a general pattern seems to emerge. The first and largest of the classes, whose members are predominantly (but not exclusively) *non-victims*, represent the contribution of the excess zeros to the distribution³¹. At the

³⁰ Although the BCS analysis optimizes one more group than the SCVS for both property crime (six compared with five) and personal crime (five compared with four), probably this is a 'degrees of freedom' phenomenon, due to differences in sample size and in numbers of variables (see Table 1). To test the effect of sample size on optimization capacity, an identical LCA model was run on sub-samples of diminishing size, each selected randomly from the overall pooled sample of BCS property crime. Results clearly confirm the importance of sample size; while twenty per cent of the overall pooled sample could support a minimum three-class solution, it would need around eighty per cent to support the optimum six-class solution, according to the ABIC.

³¹ The labels used to describe the classes in this and other solutions are merely descriptive and are employed for illustrative purposes only. As discussed earlier, the point of this analysis is to establish that the distribution is composed of ordered latent classes, rather than to identify groups. To reflect this, the following labeling

other extreme, there is a very small class of respondents with the highest probability of being frequently victimized, representing the extremely attenuated right-hand tail of the distribution, those whom we might term *Chronic Victims*. The dominant tendency, then, is for a predominance of zeros within individual crime victimization matrices. This is reflected in the ordering of the 'intermediate' classes: it is more likely for intermediate classes to have means approximating to zero on some offence types than it is for them to have means in the higher ranges, although this tendency wears off as class size diminishes. What distinguishes each intermediate class within the distribution is the dominance or absence of one or two types of crime. In sum, individual event matrices vary: those who are mainly unvictimized mostly have zeros in the cells of their event matrices, with a smattering of counts across the board. Similarly, those who are chronically victimized have many cells of their matrices with high frequency counts. In contrast, those between these two 'polar types' have different mixtures of crime types and frequencies, distinguishable by high frequencies in a few of the possible crime types.

Figure 5 about here

Figure 6 about here

Figure 7 about here

Figure 8 about here

Some differences also emerge around the general pattern: in the first place, a larger number of classes were identified for property crime compared to personal crime, presumably reflecting the larger number of crime types in the former than in the latter category (see Appendix). Second, the number of classes identified in the SCVS is consistently less than for the BCS. In part, this may reflect a difference in absolute sample size³². Nevertheless, there also seem to be qualitative differences between the Scottish and the English & Welsh distributions. Taking property crime first (Figures 5 and 6), Scotland seems somewhat more polarized than England and Wales, with a greater prevalence of both Non-Victims and Chronic Victims (and proportionately fewer Intermediate Victims). Additionally, the England and Wales (BCS) data display higher means for offences involving theft, while the Scottish (SCVS) data show higher means for criminal damage.

Regarding crimes against the person, while victims more frequently (and/or repetitively) experience the threat or actuality of violence than of theft from the person in both jurisdictions, personal victimization seem rather more widely distributed amongst the population of England and Wales than of Scotland³³.

convention is used throughout: Non-Victims > Intermediate Victims (A > B > C > D) > Chronic Victims, ordered by class size, where Non-Victims + Intermediate Victims + Chronic Victims = $\sum_{j=1}^C \pi_j$ [Equation 1].

³² See Table 1, and Fn 28.

³³ In England and Wales, 16 per cent of the population fall into the Victim classes; while the corresponding figure for Scotland is 11.6 per cent.

Nevertheless, the frequency of threatening and violent victimization is greater for Scottish than for English and Welsh victims (Figures 7 and 8). In the absence of any explanatory data, the possibility of erroneous interpretation of these differences is considerable. Furthermore, elucidating these observations would require a comparative approach that was not only grounded in theory but was able to model the partial and interaction effects of socio-structural, cultural and situational factors. Additionally, the various possible effects of selection and response bias within both samples would also need to be discounted. Nevertheless, the distinction made in this analysis between victim-prevalence and victimization-frequency seems to be a useful grammar for expressing the qualitative differences observed between the two samples.

3.2 The distribution as a mixture

Despite local variation amongst the samples, a general pattern seems to emerge amongst the sets of latent classes: mean victimization counts increase as class size diminishes, while intra-class variance between the different crime types is greater for the intermediate classes than the extreme classes. Support for the idea that this is a result of the mixing of latent processes of immunity and exposure comes from the BCS six-class solution for property crime. First of all, Figure 9 plots the cumulative proportions accounted for by each class, ordered by class size, of population and victimization. It confirms that the replication of the distribution as a set of ordered latent classes also reflects the Pareto 80/20 'power law' implicit in the skewed frequency distribution: the largest class (designated as Non-Victims) account for around 80 per cent of the population but only just over 20 per cent of the total property crime victimization, while the remaining 20 per cent of the population, in the intermediate and chronic classes, accounts for just under 80 per cent of the property crime victimization.

Figure 10 seems to confirm that this outcome is a product of a correlation of victim-prevalence and victimization-frequency. Across the ordered latent classes, mean victimization rates (expressed as a ratio of the proportion of victimization to the proportion of the population accounted for by each class) increase as the proportion of the population accounted for by each class diminishes. At one pole of the distribution is the large class of Non-Victims, indicative of the 'leverage' that the victim-prevalence component exerts over the dispersion of the distribution. Conversely, despite the smaller sizes of the 'victim classes' (especially the Chronic Class), their higher frequencies demonstrate the respective leverage of the victimization-frequency component on the distribution's dispersion. Mixing these two components together results in the distribution illustrated in Figure 9. Figure 10 resembles that expected in Figure 2, though moderated both by the large size of the victim prevalence component and the high frequencies of the victimization frequency component.

3.3 Separate categories or a continuous dimension?

Does the data generation process result in separate categories of crime victims or a continuous dimension of crime victimization risk? With reference to BCS property crime, the iteration process of the LCA towards the optimal solution suggests that the non-victim class is likely to be a stable, dominant characteristic of the distribution. From a 'starting point' of around 95 to 99% of respondents reporting zero victimization across the various property crimes (Table 1), the proportion in the class corresponding to the 'zero pole' in the Two-Class solution reduces to 88.3% but reductions in the size of this class are minimal thereafter.³⁴ There is some justification, then, for thinking that the difference between the non-victim class and the remainder of the population constitutes the most important cut-point in the distribution. Yet, it must be emphasized that this is not a purely zero group, since this class still accounts for approximately 23 per cent of the total crime victimization (Figure 9).

Put differently, this result suggests that the bulk of the classification comprises more refined classes formed mostly out of the more frequently victimized sections of the population. For instance, while the Non-victim class hardly changes at all between the Three-class and the (final) Six-class solution, the four intermediate classes together increase by over a third in population size. This process can be seen clearly in the transition from the penultimate (Five-class) to the final (Six-class solution) in Table 2. Nevertheless, even at this late stage of iteration, where there is relatively little gain in optimization (Figure 3), there is still quite a degree of re-classification going on amongst the victim classes. This suggests that there are unlikely to be other cut-points that would distinguish further stable categories beyond the basic categorization of non-victim and victim. Even so, this is a probabilistic rather than exact categorization; even in the non-victim category there remains a victimization-frequency component, although the latter becomes stronger and more characteristic of the remainder of the more frequently victimized sections of the population. As a corollary, there does not seem to be a likelihood of there being a distinctive category of multiple or high-frequency victims.

Table 2 about here

3.4 The distribution as a heterogeneous mixture of local distributions

While the previous sections have looked at the probability of population membership in each class, the next sections (3.4 and 3.5) look at the probability of class membership by level of victimization frequency, again taking the Six-class BCS property crime solution as a template. Specifically, for each victimization count reported, we use the respondent's probability of appearing in each class to calculate their estimated mean probability of appearing in each class conditional on their

³⁴ Specifically, the size of the 'non-victim' group reduces from 88.3% (2-class), 82.8% (3-class), 81.3% (4-class), 81.0% (5-class), to 79.8% (6-class).

count (overall level) of victimization. Figures 11 and 12 show the distributions of the probability of membership in the two 'polar' classes (non-victims and chronic victims) and the four intermediate classes respectively, conditional on the overall count of victimization³⁵.

Figure 11 about here

Figure 12 about here

In the first place, few counts appear to have a certainty ($p. = 1$) of falling into a particular class, other than very high and very rare frequencies. All other counts have some probability ($p.$) of being in two or more classes, although the range is most restricted for the Non-victim class. Thus, while respondents scoring zero victimization have a probability of $p. = 0.869$ of being in the Non-victim class, they also have a probability of $p. = 0.101$ of being in the Intermediate Victim-A class. In contrast, it is only the victims at the very high frequency range (a count of 14 or more) who fall exclusively into the Chronic Victim class. As such, non-negligible probabilities of being in two or more classes occur for all but the very high frequencies (i.e. of 14 or more). Since the classes sum together ([Eq. 1]), the probabilities of each count falling into each class likewise fit together. Thus, the lower probability of victims with a count of six being Chronic Victims ($p. = 0.062$) is attributable to the greater probability of victims having counts of six falling into the Intermediate A ($p. = 0.682$), Intermediate C ($p. = 0.208$), Intermediate D ($p. = 0.029$) and Intermediate B ($p. = 0.019$) victim categories respectively (Figure 12)³⁶.

The distribution of the probabilities of counts falling into each class illustrated in Figures 11 and 12 can also be interpreted as indicative of the various sub-distributions that make up the overall distribution. The differences in shape between them can likewise be seen as exemplifying the heterogeneity of the overall population. Without further information, any interpretation ventured here can only be speculative. Nevertheless, it is intriguing that the shape of the distributions of the Non-Victim and Chronic Victim classes (the latter minus the 'dips' noted above) not only most closely resemble each other (Figure 11) but also differ from the shapes of the distributions of the Intermediate Victim classes (Figure 12). The more over-

³⁵ Similar patterns were found for BCS personal crime.

³⁶ Likewise, the dip at the count of 11 in the probability of being a chronic victim ($p. = 0.642$) is attributable mainly to the probabilities of being an Intermediate C ($p. = 0.292$) or Intermediate A Victim ($p. = 0.064$); while the dip at the count of 12 ($p. = 0.486$) is attributable to the probabilities of being an Intermediate C ($p. = 0.363$), Intermediate D ($p. = 0.109$) or Intermediate A ($p. = 0.042$) Victim. These occurrences may be in part a consequence of the capping procedure employed (see Section 3.5). Specifically, respondents associated with the chronic class typically report high levels of victimization across the range of crime types, which might be subject to capping, while Intermediate Victim respondents may only experience frequent victimization in one or two types of victimization, which would be the ones that were subjected to the capping procedure.

dispersed shape of both the polar classes (Figure 11) suggests that there may be even further heterogeneity concealed within these classes: in the case of the Non-Victim class this might be attributable to a difference between those members who report low-frequency victimization and those who are consistently zero (we discuss this point in Section 4). Similarly, multiple victims can also be found outside the Chronic Victim class (see also Section 3.5) while lower frequencies can also be found within the class. In contrast, the Intermediate Classes are more homogeneous in their range of frequency-probabilities, with little suggestion of over-dispersion, albeit with varying skew shapes (Figure 12) and differences in the nature of victimization experienced.

3.5 The effect of uncapped frequencies

As discussed above (Section 2.3), extreme frequencies were capped at a maximum of 6, primarily to avoid problems of model fitting. Table 3 gives information on what might be the effect of extreme cases on the Six-Class solution for BCS property crime. In the first place, only around 0.32 per cent of respondents had their counts capped. Of these, only three respondents had more than two of their counts capped, and all were in any event classified into the Chronic Victim class. In contrast, the remaining capped frequencies were classified across the range of classes, in varying proportions. Clearly, including the uncapped extreme counts will inflate the class means, although the effect seems less for the Chronic Victim class than for the others. This provides some further assurance that the Chronic Victim class already captures the tiny minority of highly extreme frequency victims, notwithstanding the necessity of capping. Yet Table 3 also shows that high frequency victims are not necessarily the exclusive property of the Chronic Victim class. In the end, the question about how to treat the highest frequency victims may boil down to how they are to be classified: on the hand, if 'multiple' or 'repeat victims' are to be distinguished only by their higher frequency of victimization on one or more crime types, then many of them may also have overall victimization event experiences that are similar to other victims who do not have such high frequency sequences; on the other hand, if 'chronic victims' are distinguished by their common profile of high frequencies across crime types, then most of them may be multiple or repeat victims³⁷.

Table 3 about here

4. Summary, discussion and conclusions

4.1 Summary and discussion

³⁷ While negligible numbers of victims with a frequency of 14 victimizations or more were classified as anything other than Chronic Victims, below that level, substantial numbers were placed in other classes as well.

The aim of this paper was to explore the heterogeneity that might underlie the distribution of crime victimization amongst the population in order to gain some insight into its data generating process. The starting point was to think of the crime victimization experiences of each individual as comprising a 'victimizing state' (Biderman, 1981). Operationally defined, each respondent in the BCS and SCVS was considered to have their own matrix of crime victimization counts, with each cell defined by the type and frequency of victimization reported over the recall period. Various decisions about the data had to be taken in order to render these matrices suitable for analysis (Sections 2.2 and 2.3). A finite mixture methodology (employing Latent Class Analysis) was then used to generate a set of ordered latent classes that would express parsimoniously the distribution of the individual event matrices amongst sample respondents. The analysis first used the results of the LCA to replicate the characteristics of the distribution, with a view to inferring the hypothetical latent data generating processes that might be producing its over-dispersion. Second, it looked at the distribution of crime counts (as a function of the probabilities of class membership) in order to infer the presence of latent sub-population distributions within the overall distribution. The results confirmed the expectation that the distribution of individual crime victimization event matrices might be a consequence of a heterogeneous mixture of sub-population distributions that produces over-dispersion in the overall distribution.

Over-dispersion can be produced by heterogeneity within a distribution, and heterogeneity can be the multiplicative product of a mixture of processes associated with each tail of the distribution. Together they can account for the main causes of over-dispersion: both an excess of zeros, and an excess of higher frequencies, relative to the benchmark Simple Poisson model for count data (Section 1.3). Consequently, it was hypothesized that the distribution of crime victimization event matrices would be generated by a mixture of two processes: the first, producing victim-prevalence, the second, producing victimization-frequency. These processes were conceptualized as propensities for *immunity* from and *exposure* to crime victimization, respectively. The results supported the suggestion that the distribution was defined by the interaction of these two components, with the inference being that neither would account satisfactorily for the distribution independently of the other.

The findings confirm the proportionate 'power' exerted over the distribution by the 'few' high-frequency (multiple and/or repeat) victims. Nevertheless, they also demonstrate the absolute power of non-victimization over the distribution. The resulting contest between the two sources of power, as it were, gives shape to the distribution. The outcome of this analysis is that non-victimization predominates, notwithstanding the disproportionate contribution to the sum of victimization of the higher-frequency victim. Although recent attention has been paid to the excessive frequency of the distribution, relatively little attention has been paid to understanding its excessive zeroness. While crime victimization is acknowledged to be a relatively infrequent experience, affecting perhaps around twenty per cent of the population at most at any one time, the fact that victimization for the remaining eighty per cent is more or less absent has tended to be over-looked. To be sure,

there is a lengthy tradition of prevalence studies (noted in Section 1.2) but these have been 'positivistic', in the sense that they have looked at the distinction between non-victims and victims with the purpose of finding out what affects a positive (i.e. crime victimization) outcome. The suggestion of this paper is that there is also a predominant process of immunity, making it also important to understand what 'causes' a zero outcome.

Understanding the reasons for the absence of a phenomenon, or an event that does not happen, presents rather severe epistemic difficulties, in the face of which investigators tend to make the simplifying, though uncorroborated, assumption that the factors associated with 'absence' are merely the obverse of the factors that are present when the phenomenon is also present. In contrast, the method presented in this paper provides an alternative way of overcoming this problem instrumentally; that is, by conjecturing a broader class of 'predominantly but not exclusively non-victims' amongst whom the 'true zeros' are probably located. The question then becomes that of understanding what, if anything other than random chance or egregious circumstance, could disrupt the immunity that otherwise seems to pertain (Bottoms and Costello, 2009; Sparks, 1981). On the basis of this analysis, it could be argued that the class identified as Non-Victims represents the 'normal condition', since it accounts for around 80 per cent of the population³⁸. Yet this condition may not be one of absolute immunity since it still also accounts for around 20 per cent of the crime victimization. Further, within this 'normal population' there are still deviations from a norm of zero, with the distributional pattern of this deviance mirroring the heterogeneity of the distribution of the whole population, exhibiting both excessive zeroness as well as a range of frequencies (Figure 11)³⁹. In this respect, even within the 'normal condition' of the population, the propensity for immunity may still be prejudiced by occurrences of exposure to victimization (Figure 2).

The remaining 20 per cent of the population of 'predominantly but not exclusively victims' accounts for around 80 per cent of the victimization. Unlike the population of 'predominantly non-victims', it did prove possible to identify different classes on the basis of their mix of frequency and crime type. In particular, it was possible to isolate a class of chronic victims, amongst whom 'excessive' frequency cases are predominantly though not exclusively located. Not only might they be a rare type of victim but they also differ from the remainder of victims by being consistently and highly victimized across all crime types⁴⁰. Nevertheless, their rarity means they

³⁸ With the caveat that the population we are considering here is characterized only by its type and frequency of crime victimization.

³⁹ The Non-Victim class also captures substantial numbers of low-frequency victims. Victims with a frequency (f.) of 1 victimization have a probability (p.) of 0.514 of being classed in the Non-Victim category, and subsequently f. = 2, p. = 0.142, and f. = 3, p. = .03 (Figure 11).

⁴⁰ This consistency did not quite hold for the SCVS dataset (see Figures 6 and 8), though it is difficult to know whether this is due to substantive differences or artefactual differences between the surveys (see Fn. 25).

contribute relatively little to the overall total of crime victimization, even though their contribution is highly disproportionate to their prevalence (Figure 9). Additionally, many multiple or repeat victims can be found amongst the other Intermediate Victim classes (Table 3), implying that, apart from the particular sequence or series of offences they have experienced, the remainder of their victim profile is similar to other victims who may not have experienced such exceptional series. The Intermediate Victim classes can be differentiated according to their particular mix of frequency and type of crime victimization. Since they appear to be more homogeneous than the other classes (Figures 5 – 8, 12), they prompt the question of whether there might be discrete categories or syndromes of victimization within the distribution, with their own specific etiologies (a question that can be asked also of the chronic victims).

This paper cannot address this question directly, since it has employed no explanatory variables or model. Nevertheless, it can say something about the likelihood of there being discrete categories. On the one hand, the results showed a consistent pattern across the samples: a large 'Non-victim' class, a tiny Chronic Victim class, and a sizeable Intermediate Victim class, each with their different characteristic mix of crime type and frequency (Figure 9). The pattern in which these classes were formed during the iterations of the LCA was also similar between property and personal crime, and between the BCS and SCVS samples. It is therefore tempting to see these as discrete categories of crime victims with their separate syndromes of crime victimization. On the other hand, whatever the sample, the optimal number of classes always exceeded these categories, with significant shifts of membership occurring right up to the point of optimization, especially in the Victim classes (Table 2). At this finer level of optimization, there was variation in the number, size and mix of type and frequency across all the four samples. Although this might still imply that it would be possible to identify separate syndromes at a finer, more situation-specific level of analysis, an explanation would need to balance the generality of the three-fold classification against the specificity of the more local solutions.

Perhaps a more compelling reason why this question cannot be resolved empirically is that it begs another question that is ultimately epistemological: whether the data generating process is fundamentally categorical or dimensional in character? The approach adopted in this paper has been to conjecture that the latent distribution of crime victimization is a multiplicative product of the mixing of two theoretical propensities (for immunity from, and exposure to, crime victimization) that produce a heterogeneous mixture of victim classes. As such, the distribution has a dual nature: its product might consist of categories, while its generative process might consist of dimensions. Thus, it might be argued, whether categories or dimensions are regarded as of most interest depends on whether the concern is forensic and diagnostic, favoring a taxonomic approach that seeks to identify precise probabilities for individual cases; or whether it is epidemiological and prognostic, favoring an understanding of process and development that seeks to identify general probabilities for populations (Pickles and Angold, 2003). Analogous to the proven

scientific usefulness of a dual conception of light, the 'duality' of the phenomenon in question offers to expand the possibilities of enquiry (see also Maltz, 2009)⁴¹.

4.2 Conclusions

This paper has been concerned with some foundational issues that contribute towards understanding the distribution of a physical phenomenon; that is, the occurrence of criminally victimizing events as they occur to their victims. Paradoxically, our understanding of the distribution of crime victimization seems to have become more confused the more that empirical research has progressed (Section 1.2). One reason for the confusion that has arisen stems from the dual nature of the frequency distribution itself. A frequency distribution is a way of arranging count-data on two dimensions: in this case, a count of crime victims, and a count of crime victimization events, within an implicit third, temporal, dimension. Naturally, this has given rise to two analytic approaches: one that examines crime victims in terms of the victimization they experience (similar to studying the differing susceptibilities of the potential hosts of a disease); the other that examines how crime victimization (through the agency of its perpetrators) selects its victims (similar to studying the contagiousness of the disease itself). And as each research approach has evolved, so it has become ever harder to reconcile them, even though they derive from data with a common structure.

Despite their many measurement limitations, crime victimization surveys remain a primary source for studying the occurrence of crime victimization events as observed from the self-reports of crime victims (Miethe and Meier, 1994). For the past thirty years, the Routine Activity 'theorem', in combination with the 'lifestyle-exposure' approach, has served as the baseline model for explaining the occurrence of crime victimization events.⁴² Initially, empirical research on data from crime victimization surveys conceptualized this as a *discrete-outcome* model, concerned with predicting the likelihood of experiencing a victimization event, as in Cohen and Felson's (1979) initial formulation⁴³. The baseline model was further extended to predict the

⁴¹ Nevertheless, it should be remembered that this is a methodological precept. As Karl Popper remarked "...we must always clearly distinguish between an 'essentialist' explanation with appeals to the nature of things and a 'descriptive' explanation which appeals to a Law of Nature, i.e. to the description of an observed regularity. Of these two kinds of explanation only the latter is admissible in physical science" (1969: 169).

⁴² See Cohen and Felson (1979) and Hindelang, et al. (1978) respectively, and for further elaboration, *inter alia*, Cohen et al. (1981), Miethe and Meier (1994) and Osgood et al. (1996).

⁴³ That is, "...the probability that a violation will occur at any specific time and place might be taken as a function of the convergence of likely offenders and suitable targets in the absence of capable guardians" (Cohen and Felson, 1979: 590)

likelihood of 'repeat victimization', dependent on the occurrence of 'initial victimization' (Farrell et al., 1995). Nevertheless, the resulting *double-hurdle* conception that has developed by default (Osborn et al., 1995) has been an uneasy hybrid, throwing-up many empirical puzzles and difficulties of model specification (some of which are discussed in Section 1.3), while resulting in poor explanatory performance that has not much improved over the past twenty years since Meier and Miethe's (1993) appraisal (Hope, 2007a).

The current state of play makes it impossible to work out whether the fault lies with the baseline theoretical model or with the manner in which it has been operationalized in empirical research. This impasse may be due to a fundamental difficulty with the way in which the baseline model has been conceptualized, namely, that it has been deployed in order to predict exact probabilities for the occurrence of singular events of crime victimization. This approach ushers-in a host of epistemic difficulties when dealing with observed data (Gilles, 2000; Popper, 1959), not least of which is the difficulty of predicting the counterfactual non-occurrence of events, including the assumption that the non-occurrence of a (crime victimization) event is merely the obverse of its positive (observed) occurrence (which is a version of the 'omitted variable' problem, see Lütkepohl, 1982). Yet the present analysis has suggested that the distribution of criminal victimization events as they occur to victims cannot be properly defined without consideration of the role played by non-occurrence, that is, by 'zero-inflation'.

A challenge that could be addressed to this paper is: 'given what we already know about the nature of victimization and its distribution, what more do these analyses tell us?'⁴⁴. Our answer is that this is a difficult question to address at present, precisely because our approach calls into question the basis of current knowledge. In as much as empirical research has more or less ignored the effect of the zeros that dominate the distribution (or where they have not been ignored, they have been discounted), the challenge of whether or not zero-inflation matters cannot be addressed until a method of analysis is deployed that incorporates both the influence of non-occurrence (what we have called 'immunity') as well as that of occurrence (what we have called 'exposure'). This lies not so much in finding new data but more in modeling existing data in a way that would allow new substantive evaluation of what we think we already know (Hope and Trickett, 2008).

The two-class solution developed by our analysis, illustrated in Figure 13 (for BCS property crime), demonstrates the potential of such an approach. The simple premise of the two-class solution is that occurrence and non-occurrence can be conceptualized as separate probability distributions of membership in one of two classes, 'victims' and 'non-victims'⁴⁵. In other words, the two-class solution reconstructs the observed one-dimensional frequency distribution (also illustrated in

⁴⁴ We are grateful to one of our anonymous reviewers for posing this question.

⁴⁵ Where measurement is by positive integers, a two-group solution is the inverse of the frequency distribution, replicating the cut-point between all positive values and zero.

Figure 13) as a hypothetical world consisting of the two dimensions – the propensities of immunity and exposure - modeled by the observed probability distributions of being a victim and of being a non-victim. The validity of this conceptualization is supported by the similarity between the probability distributions generated from the two-class solution (Figure 13) and the frequency distributions of victim-prevalence and victimization-frequency generated from the ordered latent classes of the optimum (six-class) solution (Figure 10).

Figure 13 about here

In this new, hypothetical baseline model (premised on the Two-Crossings Theorem discussed in Section 1.3), the two probability functions would appear to ‘cross’ at the point that allocates exactly half the sample to each class and gives one ‘quantum’ of victimization to each (Figure 13). This might then be interpreted as a model of a hypothetical world where the two propensities of immunity and exposure do not mix together, and so do not generate the over-dispersed frequency distribution evident in observed data. In contrast, the solutions optimized by subsequent iterations of the LCA represent an observed world in which the probability functions (propensities) do mix together to produce empirical outcomes, whose probabilities are estimated from the latent classes. The mixing thus creates a single probability distribution ranging from a zero-probability of victimization at one pole to a unitary-probability at the other, characterized both by zero-inflation and high-frequency. Consequently, this distribution is latently heterogeneous and empirically over-dispersed. As a result of the mixing of victimization propensities, the bulk of crime victimization events fall intermediately between these two polar extremes producing a set of possibly discrete but homogeneously distributed (i.e. non-dispersed) outcomes. Finally, a residuum of victimization (at very low frequency) is allocated to those who gravitate towards the zero-pole of the population distribution, while a residuum of victims, victimized at very high frequency, remains at the unitary-pole. It turns out that there is a far greater tendency amongst the general population to gravitate towards the zero-pole than towards the unitary pole of victimization.

The analyses presented here refer to ‘general populations’. There may be also members of other ‘special’ populations – for instance, children, prisoners, criminal networks and gangs, victims of corporate crime, victims of domestic violence, the homeless, etc. - that might have different crime victimization distributions, and who thus ‘deviate’ from the general population in the nature of their victimizing state. Analogous to deviancy theories of offending, crime victimization may also have a quality of deviance (that is, departing from the social norm of non-victimization from direct-contact crime) that would place a special focus substantively on the propensity for victimization-exposure in preference to that of victim-prevalence. For instance, Bottoms and Costello (2009) identify a relatively high-level of sustained repeat victimization amongst households containing at least one member who had previously been known to the police, usually as an offender. In this respect, the exposure process might stem from the nexus between offending and victimization (Lauritsen and Laub, 2007). The present model may not be sensitive enough to the particular circumstances of certain crime types. The survey-derived distribution of

domestic violence incidents, for example, albeit displaying a similar property of zero-inflation, would seem to exhibit a much 'thicker' right-tail than that of other types of crime (Hope, 2007a), presumably reflecting the relational and escalatory nature of intimate partner violence where, once initiated, repeat victimization rapidly becomes the norm.

Even so, regardless of whether the population is general or special in some way, the hypothetical baseline model may have a general, heuristic value in two respects: first, in dealing with two probability distributions, instead of one frequency-distribution, it is possible to assess for any particular distribution of crime victimization, the relative weight in explanation to be placed on each tail. As such, the suggested baseline model may make some tentative claim to being a general model, applicable to all distributions of crime victimization events. Further, since the baseline model appears to apply to both property crime and personal crime victimization, it appears to overcome a major stumbling block of previous theoretical models in promising to provide a common ground on which to develop integrated theories applicable to a broad range of crime types (Miethe and Meier, 1993).

A second heuristic value of the proposed, hypothetical baseline model is to avoid the trap of over-determination into which efforts to predict singular events often fall. For instance, we cannot assess the extent to which any particular observed distribution of crime victimization events represents the totality of all possible conjunctions of the conditions identified by the Routine-activity/Lifestyle model net of our capacity to observe that totality. In contrast, the heuristic value of the hypothetical baseline model derives from the comparison between prior expectations - for instance, the base rate probabilities derived from the 'two-class' solution (discussed above) - and posterior probabilities derived from observed frequency distributions - for instance, the conditional probabilities estimated from the set of ordered latent classes (see Kahneman and Tversky, 1979). A consequence of avoiding over-determination lies in reinforcing a salutary message: that complete immunity may be as unlikely as total exposure; apparently 'immune' individuals may sometimes be victimized while even the most chronic victims may experience some periodic cessation from their suffering.

The present baseline model extends the seminal approach taken by Trickett et al. (1992), which likewise defined the crime incidence rate as a multiplicative product of the prevalence rate and a rate of victimization frequency, termed 'vulnerability' or 'concentration'.⁴⁶ In turn, this analysis stimulated interest in the possibility that concentration (or victimization-frequency) might be manipulated independently of victim-prevalence in order to bring about changes in the incidence of crime events (Farrell, 1995). Nevertheless, the trend in the crime victimization (incidence) rate in England and Wales (measured by the BCS), suggests it has been more susceptible to

⁴⁶ If the number of victims = V, the number in the population = P, and the number of victimization events = C, then Prevalence = V/P , Concentration = C/V , and Incidence = $C/P = (V/P)(C/V)$ (Hope, 2007b: 103).

changes in victim-prevalence than changes in the crime-victimization frequency (Hope, 2007b). Yet, does this mean that there are fewer susceptible victims in the population (possibly indicating an increase in the prevalence of immunity) or less frequent crime-victimization (possibly indicating a decline in the frequency of exposure); and how much weight should we place on each possibility? The results from this present analysis (like those from Trickett et al., 1992) suggest that this is a difficult question to answer descriptively since the crime rate (or the crime victimization frequency distribution) is, fundamentally, a simultaneous, multiplicative product of both these components together.

Considering both processes of immunity and exposure simultaneously might generate models that may be more accurate representations of the crime victimization experiences of the general population. This may require analytic approaches, including experimentation and simulation, which can manipulate the separate components of victim-prevalence and victimization-frequency hypothetically in order to uncover the latent data generating processes that are embedded within the observable behavior of aggregate crime rates. It is to be hoped that the present analysis is an encouraging step in that direction.

Appendix: Definitions of Property and Personal Crime

A1. Questions Used to Measure Property Crime (taken from the Scottish Crime Victimization Survey 2006)

Since [1st April 2005], (apart from anything you have already mentioned,) has anyone GOT INTO your home without permission and STOLEN or TRIED TO STEAL anything?

(Apart from anything you have already mentioned,) in that time did anyone GET INTO your home without permission and CAUSE DAMAGE?

(Apart from anything you have already mentioned) in that time have you had any evidence that someone has TRIED to get in without permission to STEAL or to CAUSE DAMAGE?

(Apart from anything you have already mentioned), in that time was anything (else) stolen OUT OF your home? For example, by a guest, a workman or anyone else there with your permission.

And (apart from anything you have already mentioned), in that time was anything (else) that belonged to someone in your household stolen FROM OUTSIDE YOUR HOME – for example, from the doorstep, the garden, a shed, outhouse or garage? Please don't include milk bottles or newspapers.

And again, (apart from anything you have already mentioned), in that time has anyone deliberately DAMAGED or DEFACED your home or anything outside it (APART FROM A MOTOR VEHICLE) that belonged to YOU or ANYONE ELSE in your household?

A2. Questions Used to Measure Personal Crime (taken from the Scottish Crime Victimization Survey 2006)

Since [1st April 2005], (apart from anything you have already mentioned), was anything you were CARRYING STOLEN out of your hands or from your pockets or from a bag or case you were carrying?

(Apart from anything you have already mentioned), in that time has anyone TRIED to STEAL something you were carrying out of your hands or from your pockets or from a bag or case you were carrying?

And again, (apart from anything you have already mentioned), since [1st April 2005] has anyone, including people you know well, DELIBERATELY HIT YOU with their fists, or with a weapon of any sort, or kicked you, or used force or violence on you in any other way?

Heterogeneity in the Frequency Distribution of Crime Victimization

And (apart from anything you have already mentioned), in that time, has anyone THREATENED to damage things of yours or THREATENED to use force or violence on you in any way that actually frightened you?

References

Allen, J. (2007) Survey Assessments of Police Performance in the British Crime Survey. In Hough, M. And Maxfield, M. *Surveying Crime in the 21st Century*. Monsey, NY: Criminal Justice Press

Biderman, A.D. (1981). Sources of data for victimology. *Journal of Criminal Law and Criminology*, 72: 789-817.

Bolling, K., Grant, C. and Sinclair, P. (2007). *British Crime Survey 2006/07 Technical Report Volume 2*. UK Data Archive: Study Number 5755.

Bottoms, A., and Costello A. (2009). Crime Prevention and the Understanding of Repeat Victimization: A Longitudinal Study. In P. Knepper, P., Doak, J. and Shapland, J. (eds.) *Urban Crime Prevention, Surveillance, and Restorative Justice*. Boca Raton: Taylor and Francis.

Brown, M. (2007). *The 2006 Scottish Crime and Victimization Survey 2006 Technical Report (Volume 1)* UK Data Archive: Study Number 5784.

Brame, R., Nagin, D.S. Wasserman, L. (2006). Exploring Some Analytical Characteristics of Finite Mixture Modeling. *Journal of Quantitative Criminology*. 22: 31--59.

Cameron, A.C. and Trivedi. P.K. (1998). *Regression Analysis of Count Data*. Cambridge University Press, Cambridge, UK.

Cantor, D. and Lynch, J.P. (2007). Addressing the challenge of costs and error in victimization surveys: the potential of new technologies and methods. In Hough, M. and Maxfield, M. (Eds.) *Surveying Crime in the 21st Century: commemorating the 25th Anniversary of the British Crime Survey*. Crime Prevention Studies Vol. 22. Monsey, NY: Criminal Justice Press.

Cohen, L.E. and M. Felson (1979). Social change and crime rate trends: a routine activity approach. *American Sociological Review*, 44: 588-608.

Cohen, L. E., J.R. Kluegel and K.C. Land (1981). Social inequality and predatory criminal victimization: an exposition and test of a formal theory. *American Sociological Review*, 46: 505-524.

Eck, J.E., R.V. Clarke and R.T. Guerette (2007). Risky facilities: crime concentration in homogeneous sets of establishments and facilities'. *Crime Prevention Studies*, 21: 225-264.

Ellingworth, D., Hope, T., Osborn, D.R., Trickett, A. and Pease, K. (1997) Prior Victimization and Crime Risk. *International Journal of Risk, Security and Crime Prevention*. 2: 201--214

Ellingworth, D. G. Farrell and K. Pease (1995). A Victim is a Victim is a Victim – chronic victimization in four sweeps of the British Crime Survey. *British Journal of Criminology*, 35: 360–365.

Farrell, G. (1995). Preventing Repeat Victimization. In Tonry, M. and Farrington, D.P. (Eds.) *Building a Safer Society: Strategic Approaches to Crime Prevention*. Crime and Justice, Vol. 19. Chicago: University of Chicago Press.

Farrell, G. and Pease, K. (2007). The sting in the tail of the British Crime Survey: multiple victimizations. In Hough, M. and Maxfield, M. (Eds.) *Surveying Crime in the 21st Century: commemorating the 25th Anniversary of the British Crime Survey*. Crime Prevention Studies Vol. 22. Monsey, NY: Criminal Justice Press.

Farrell, G., Phillips, C. and Pease, K. (1995). Like taking candy: why does repeat victimization occur? *British Journal of Criminology*, 35: 384-399.

Farrell, G. and Pease, K. (1993). *Once Bitten, Twice Bitten: Repeat Victimization and its Implications for Crime Prevention*. Crime Prevention Unit Paper 46. London: Home Office.

Gillies, D. (2000). *Philosophical Theories of Probability*. London: Routledge .

Hacking, I. (2001). *An Introduction to Probability and Inductive Logic*. Cambridge: Cambridge University Press.

Hindelang, M.S., Gottfredson, M. and J. Garofalo, J. (1978). *Victims of Violent Crime*. Ballinger, Cambridge, MA.

Hope, T. (2001). Crime Victimization and Inequality in Risk Society . In Matthews, R. and Pitts, J. (eds.) *Crime Prevention, Disorder and Community Safety*. London: Routledge.

Hope, T. (2007a). Theory and Method: The Social Epidemiology of Crime Victims. In S. Walklate (Ed.) *Handbook on Victims and Victimology*. Cullompton: Willan

Hope, T. (2007b). The distribution of household property crime victimization : insights from the British Crime Survey. In Hough, M. and Maxfield, M. (Eds.) *Surveying Crime in the 21st Century: commemorating the 25th Anniversary of the British Crime Survey*. Crime Prevention Studies Vol. 22. Monsey, NY: Criminal Justice Press.

Hope, T. and Trickett, A. (2004). La Distribution de la Victimation dans la Population, *Déviance et Société*, 28 : 385--404.

Hope, T. and Trickett, A. (2008). The Distribution of Crime Victimization in the Population . *International Review of Victimology*, 15: 37--58.

Heterogeneity in the Frequency Distribution of Crime Victimization

Hope, T., Bryan, J. Osborn, D. and Trickett, A. (2001). The Phenomena of Multiple Victimization: The Relationship Between Personal and Property Crime Risk. *British Journal of Criminology*, 41: 595--617.

Howson, C. and Urbach, P. (1989). *Scientific Reasoning: the Bayesian approach*. La Salle, IL: Open Court.

Kahneman, D. and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica*, 47: 263-91.

Karlis, D. And E. Xekalaki (2005). 'Mixed Poisson Distributions' *International Statistical Review*, 73: 35-58.

Lauritsen, J.L. and Davis Quinet, K.F. (1995). Repeat Victimization Among Adolescents and Young Adults. *Journal of Quantitative Criminology*, 11: 143-166.

Lauritsen, J.L and Laub, J.H. (2007). Understanding the link between victimization and offending: new reflections on an old idea. In Hough, M. and Maxfield, M. (Eds.) *Surveying Crime in the 21st Century: commemorating the 25th Anniversary of the British Crime Survey*. Crime Prevention Studies Vol. 22. Monsey, NY: Criminal Justice Press.

Lütkepohl, H. (1982). Non-causality due to omitted variables. *Journal of Econometrics*, 19: 367-378.

McCutcheon, A. (1987). *Latent Class Analysis*. London: Sage Publications

McDowell, D. (2010). 'The Present and Possible Future of Quantitative Criminology'. *Journal of Quantitative Criminology*, 26: 429—435.

McLachlan, G and Peel, D (2000). *Finite Mixture Models*. New York: Wiley

Magidson, J. and Vermunt, J.K. (2004). Latent Class Models . in Kaplan, D. (Ed.), *The Sage Handbook for Quantitative Methodology*. Thousand Oaks, CA: Sage Publications, pp. 175--198.

Maltz, M. (2009). Waves, particles and crime. In Weisburd, D., Bernasco, W., and Bruinsma, G.J.N. (Eds.) *Putting Crime in its Place: units of analysis in geographic criminology*. New York: Springer Science+Business Media.

Meier, R.F. and T.D. Miethe (1993). Understanding theories of criminal victimization. *Crime and Justice*, 17: 459-499.

Miethe, T. D and R. F. Meier (1994). *Crime and Its Social Context*. Albany, NY: SUNY Press

Heterogeneity in the Frequency Distribution of Crime Victimization

Morgan, F. (2007) Initial and Repeated Burglary Victimization: Victim Vulnerability, Same Offender Involvement and Implications for Theory and Crime Prevention. Unpublished PhD thesis, Perth: University of Western Australia.

Mullahy, J. (1997). 'Heterogeneity, excess zeros and the structure of count data models'. *Journal of Applied Econometrics*, 12: 337-350.

Muthen, B. (2001). Latent Variable Mixture Modeling . in Marcoulides, G and Schumacker, R, *New Developments and Techniques in Structural Equation Modeling*. New Jersey: Lawrence Erlbaum

Muthen, L. and Muthen, B. (1998-2008). *MPlus User s Guide*. Fifth Edition. Los Angeles: Muthen and Muthen

Nagin, D.S. and Tremblay, R.E. (2005). Developmental Trajectory Groups: Fact or Useful Statistical Fiction? *Criminology*, 45: 873-904.

Nelson, J. F. (1980). Multiple Victimization in American Cities: A Statistical Analysis of Rare Events. *American Journal of Sociology*, 85: 870--891.

Nylund, K., Asparouhov, T. and Muthen, B. (2007). Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study *Structural Equation Modeling*, 14: 535—569

Osborn, D.R. and A. Tseloni (1998). The Distribution of Household Property Crimes. *Journal of Quantitative Criminology*, 14: 307-330.

Osborn, D.R., Trickett, A., Elder, R. (1992). Area characteristics and regional variates as determinants of area property crime levels. *Journal of Quantitative Criminology*, 8: 265-285.

Osborn, D.R., Ellingworth, D., Hope, T., and Trickett, A. (1996) Are Multiply Victimized Households Different? *Journal of Quantitative Criminology*, 12: 223-245.

Osgood, D.W, J. K.Wilson, P. M. O'Malley, J. G. Bachman and L. D. Johnston (1996). Routine Activities and Individual Deviant Behavior. *American Sociological Review*, 61: 635-655

Pearl, J. (2000). *Causality: Models, Reasoning and Inference*. Cambridge: Cambridge University Press.

Pease, K. (1998). Repeat Victimization: Taking Stock. *Crime Detection and Prevention Series Paper 90*. London: Home Office.

Pease, K. and Farrell, G (2007). Repeat victimization. In R. Wortley, and L. Mazerolle (Eds.) *Environmental Criminology and Crime Analysis*. Cullompton: Willan

Heterogeneity in the Frequency Distribution of Crime Victimization

Pickles, A. and Angold, A. (2003). Natural Categories or Fundamental Dimensions: On Carving Nature at the Joints and the Rearticulation of Psychopathology . *Development and Psychopathology*, 15: 529--551.

Pitcher, A. B. and Johnson, S. D. (2011). Exploring Theories of Victimization Using a Mathematical Model of Burglary. *Journal of Research in Crime and Delinquency*, 48: 83-109

Planty, M. and Strom, K. J. (2007). Understanding the role of repeat victims in the production of annual US victimization rates. *Journal of Quantitative Criminology*, 23: 179-200.

Popper, K. R. (1983, 1957). *Propensities, Probabilities and the Quantum Theory*. In D. Miller (Ed.). *A Pocket Popper*. Oxford: Fontana Paperbacks.

Popper, K.R. (1969). *Conjectures and Refutations: the growth of scientific knowledge*. Third Edition (revised). London: Routledge and Kegan Paul.

Popper, K.R. (1959). *The logic of Scientific Discovery*. London: Hutchinson.

Sampson, R.J. and Laub, J.H. (2005). Seductions of method: rejoinder to Nagin and Tremblay's 'Developmental trajectory groups: fact or fiction?' *Criminology*, 43: 905-913.

Sampson, R.J. and Laub, J.H. (2003). Life course desisters? Trajectories of crime among delinquent boys followed to age 70. *Criminology*, 41: 555-592.

Schwarz, G (1978) Estimating the Dimension of a Model . *Annals of Statistics*, 6: 461--464.

Sherman, L.W. (2007). The Power Few: Experimental Criminology and the Reduction of Harm: the 2006 Joan McCord Prize Lecture. *Journal of Experimental Criminology*, 3: 299-321.

Short, M.B., D'Orsogna, M. R., Brantingham, P. J. and Tita, G. E. (2009). Measuring and Modeling Repeat and Near-Repeat Burglary Effects. *Journal of Quantitative Criminology*, 25: 325-339.

Sparks, R. F. (1981). Multiple Victimization: Evidence, Theory and Future Research. *Journal of Criminal Law and Criminology*. 72: 762--788.

Sparks, R.F., Genn, H., and Dodd, D. (1977). *Surveying Victims*. Wiley: London.

Heterogeneity in the Frequency Distribution of Crime Victimization

Tipping, S., Hussey, D., Wood, M. and Hales, J. (2010). *British Crime Survey: Methods Review 2009 Final Report*. London: National Centre for Social Research (NatCen).

Trickett, A., Osborn, D.R., Seymour, J., and Pease, K. (1992). What is Different About High Crime Areas? *British Journal of Criminology*, 32: 81--90.

Tseloni, A. (2006). Multi-level Modeling of the Number of Property Crimes: Household and Area Effects. *Journal of the Royal Statistical Society, Series A*, 169: 205--233.

Tseloni, A. and Pease, K. (2003). Repeat Victimization: 'Boosts' or 'Flags'? *British Journal of Criminology*, 43: 196—212.

Tseloni, A. and Pease, K. (2004). Repeat Personal Victimization: Random Effects, Event Dependence and Unexplained Heterogeneity. *British Journal of Criminology*, 44: 931—945.

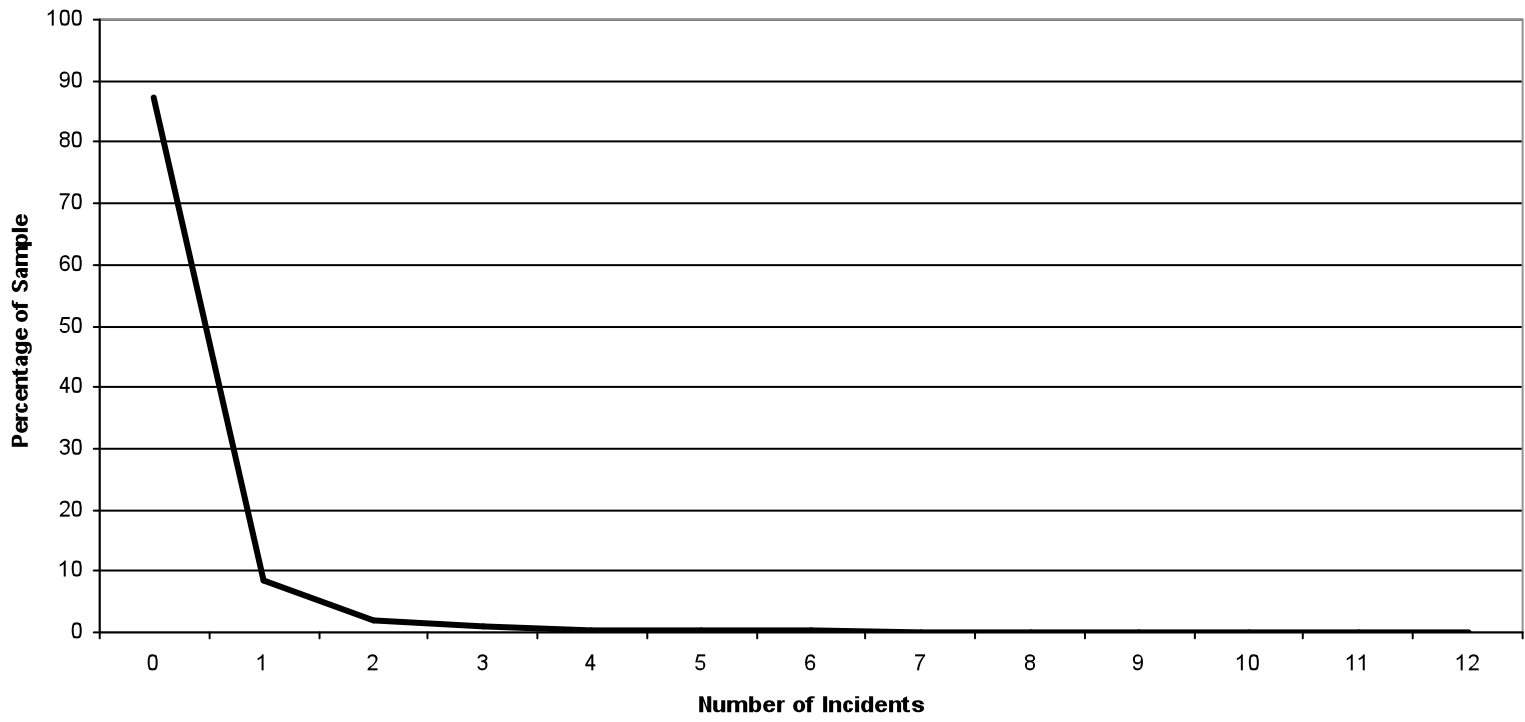
Tseloni, A., Ntzoufras, I., Nicolaou, A. and Pease, K. (2010). Concentration of personal and household crimes in England and Wales. *European Journal of Applied Mathematics*, 21: 325-348.

Tseloni, A., Osborn, D.R., Trickett, A. and Pease, K. (2002). Modeling Property Crime Using the British Crime Survey: What Have we Learnt? *British Journal of Criminology*, 42: 109--128.

Wittebrood, K. and P. Nieuwebeerta. (2000). Criminal Victimization During One's Life Course: The Effects of Previous Victimization and Patterns of Routine Activities. *Journal of Research in Crime and Delinquency*, 37: 91--122.

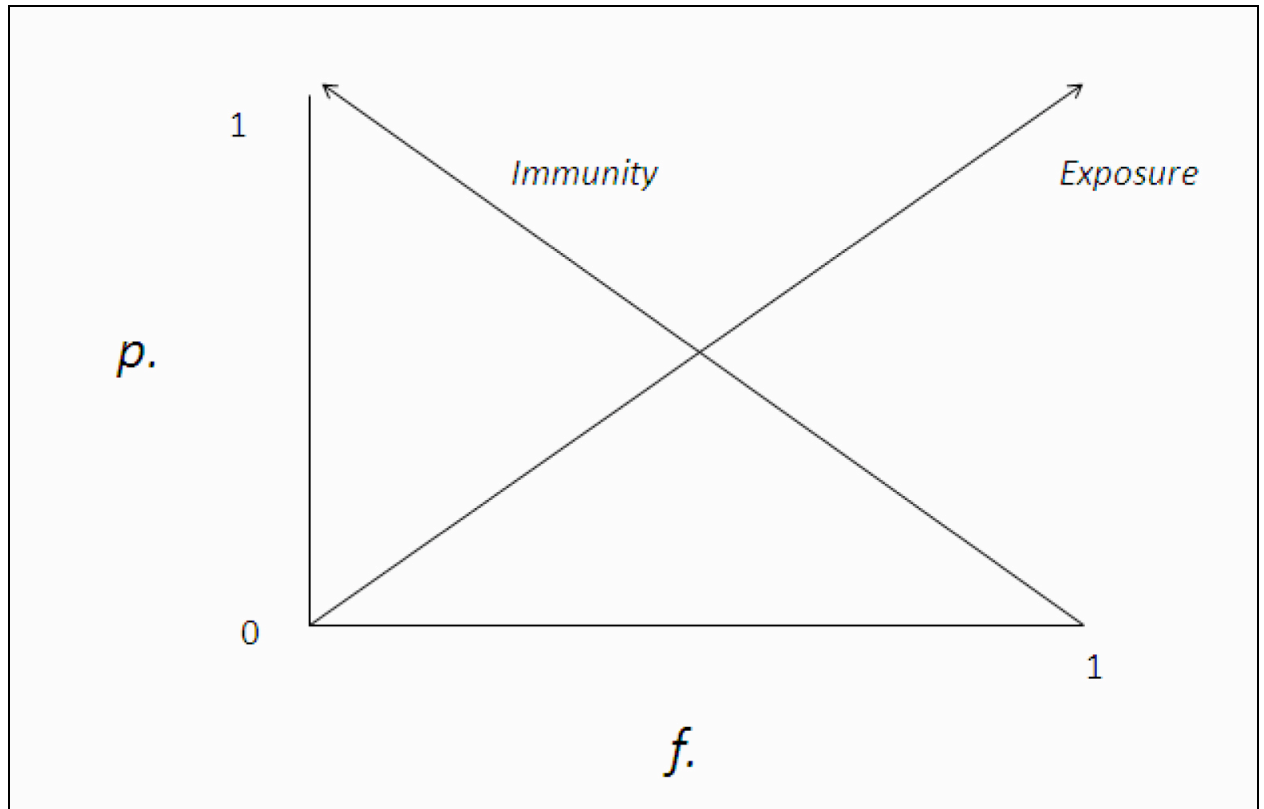
Yang, C (2006). Evaluating Latent Class Analyses in Qualitative Phenotype Identification, *Computational Statistics & Data Analysis*, 50: 1090--1104.

Figure 1 Distribution of Property Crime Victimization (SCVS 1993-2006)



Unweighted SCVS Sample: 1993 - 5030, 1996 - 5045, 2000 - 5059, 2003 - 5041, 2006 - 4988, Total - 25163

Figure 2: Probability ($p.$) and Frequency ($f.$) of Crime Victimization



Heterogeneity in the Frequency Distribution of Crime Victimization

Table 1: Descriptive Statistics for Samples Used in Latent Class Models of Property and Personal Victimization

Type of Victimization	Combined BCS Sample					Combined SCVS Sample				
	Percentage Zeros	Mean	Std. Dev	Variance	Ratio of Variance to Mean	Percentage Zeros	Mean	Std. Dev	Variance	Ratio of Variance to Mean
Property Crime										
Defaced Property (Outside)	95.1	0.09	0.05	0.25	2.72	95.5	0.09	0.51	0.26	3.06
Stolen Property (Outside)	94.6	0.08	0.40	0.16	1.97	95.6	0.06	0.33	0.11	1.82
Property Stolen from Home	99.3	0.01	0.14	0.02	2.30	98.8	0.02	0.20	0.04	2.35
Tried to Gain Entry to Commit Theft/Damage	97.2	0.04	0.26	0.07	1.86	97.6	0.03	0.23	0.05	1.76
Entered Property and Caused Damage	99.7	0.00	0.09	0.01	2.16	99.7	0.00	0.09	0.01	2.12
Entered Property and Committed Theft	97.2	0.03	0.23	0.05	1.49	98.2	0.02	0.18	0.03	1.46
Personal Crime										
Property Stolen from Person	97.8	0.02	0.18	0.03	1.39	98.6	0.01	0.14	0.02	1.34
Attempted Theft of Property from Person	99.1	0.01	0.13	0.02	1.58	99.3	0.01	0.14	0.02	2.00
Experienced Violence	96.4	0.06	0.42	0.18	2.87	96.4	0.06	0.44	0.19	2.97
Received Threat of Violence	95.2	0.10	0.57	0.32	3.29	95.7	0.09	0.56	0.31	3.43
	Unweighted BCS Sample: 1992 - 11713, 1996 - 16348, 2001 - 8927, 2003/04 -37931, 2006/07 - 47027, Total - 121946					Unweighted SCVS Sample: 1993 - 5030, 1996 -5045, 2000 - 5059, 2003 - 5041, 2006 - 4988, Total – 25163				

Figure 3: ABIC Statistics for Models of Property Crime, BCS and SCVS Datasets

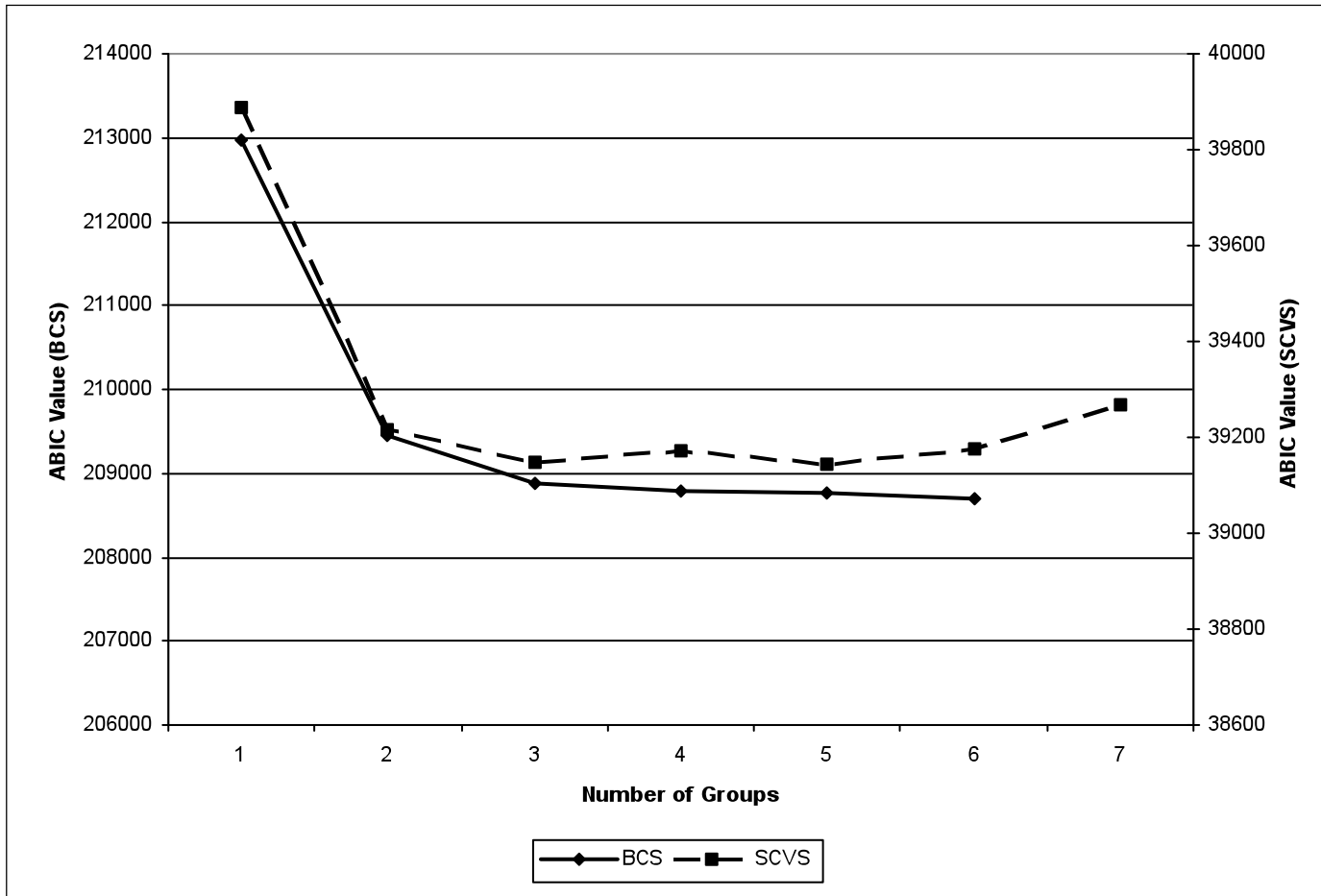


Figure 4: ABIC Statistics for Models of Personal Crime, BCS and SCVS Datasets

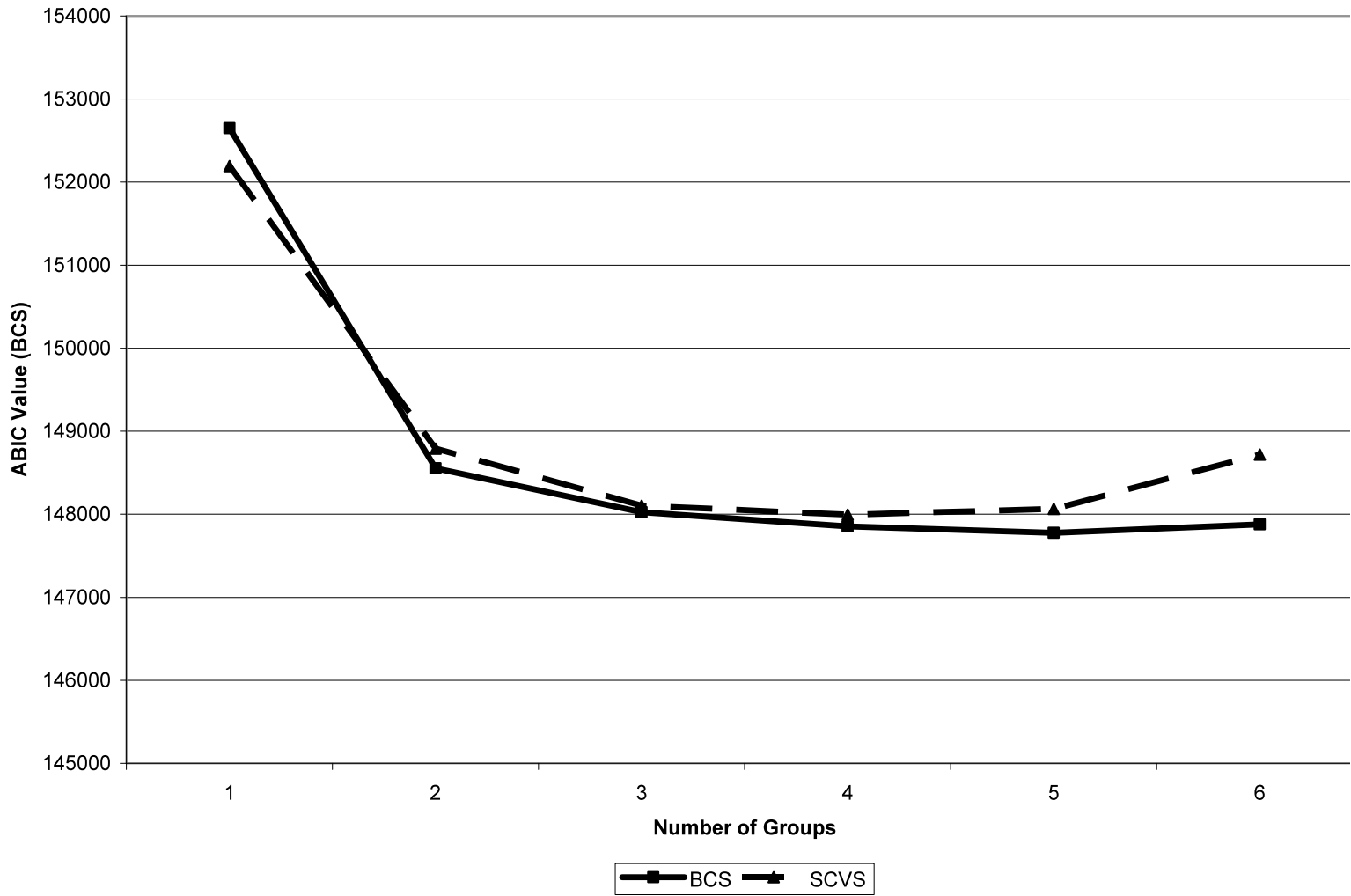


Figure 5: Mean Number of Victimizations by Property Crime Type – Six Class BCS Model

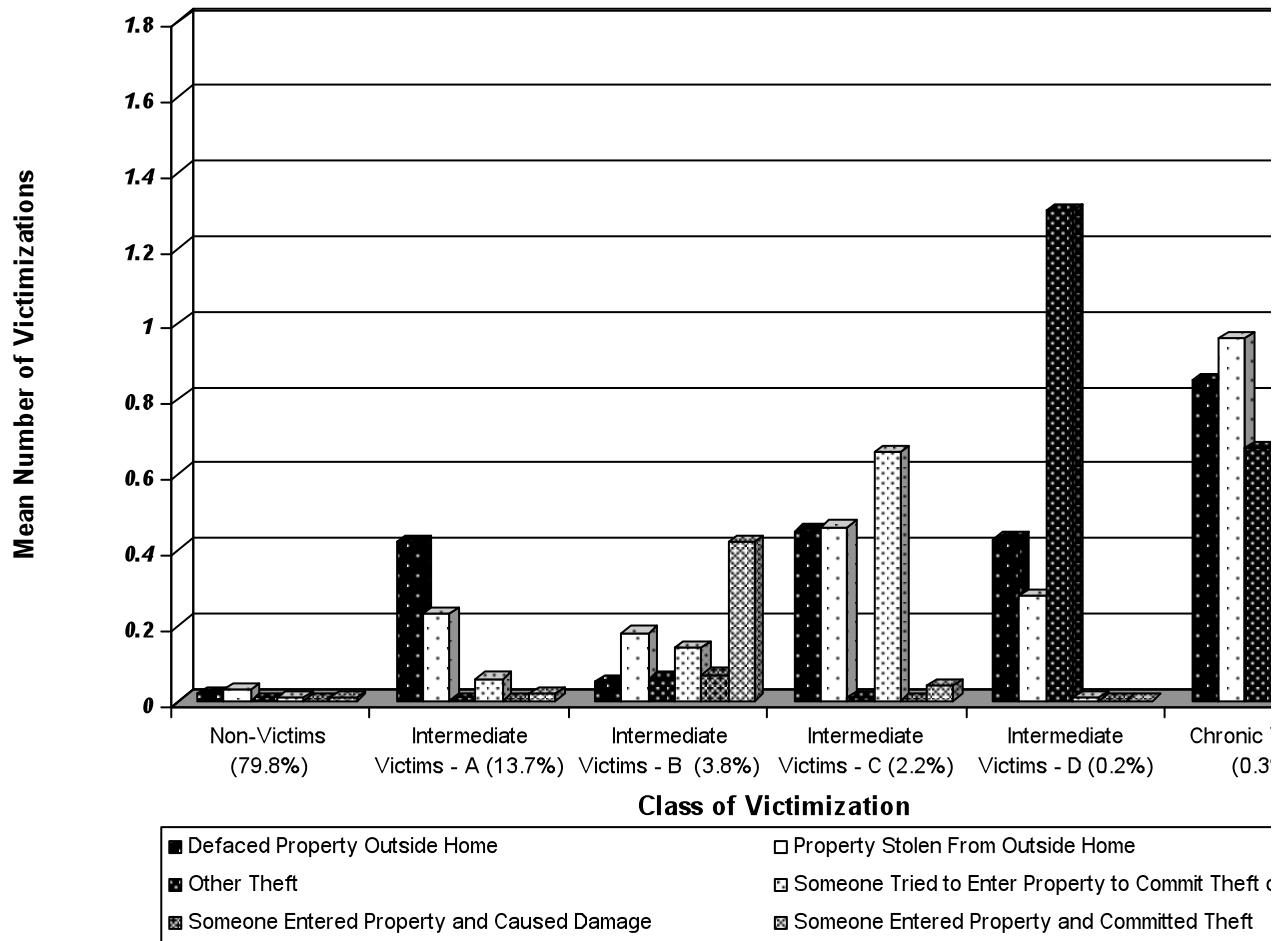


Figure 6: Mean Number of Victimizations by Property Crime Type - Five Group SCVS Model

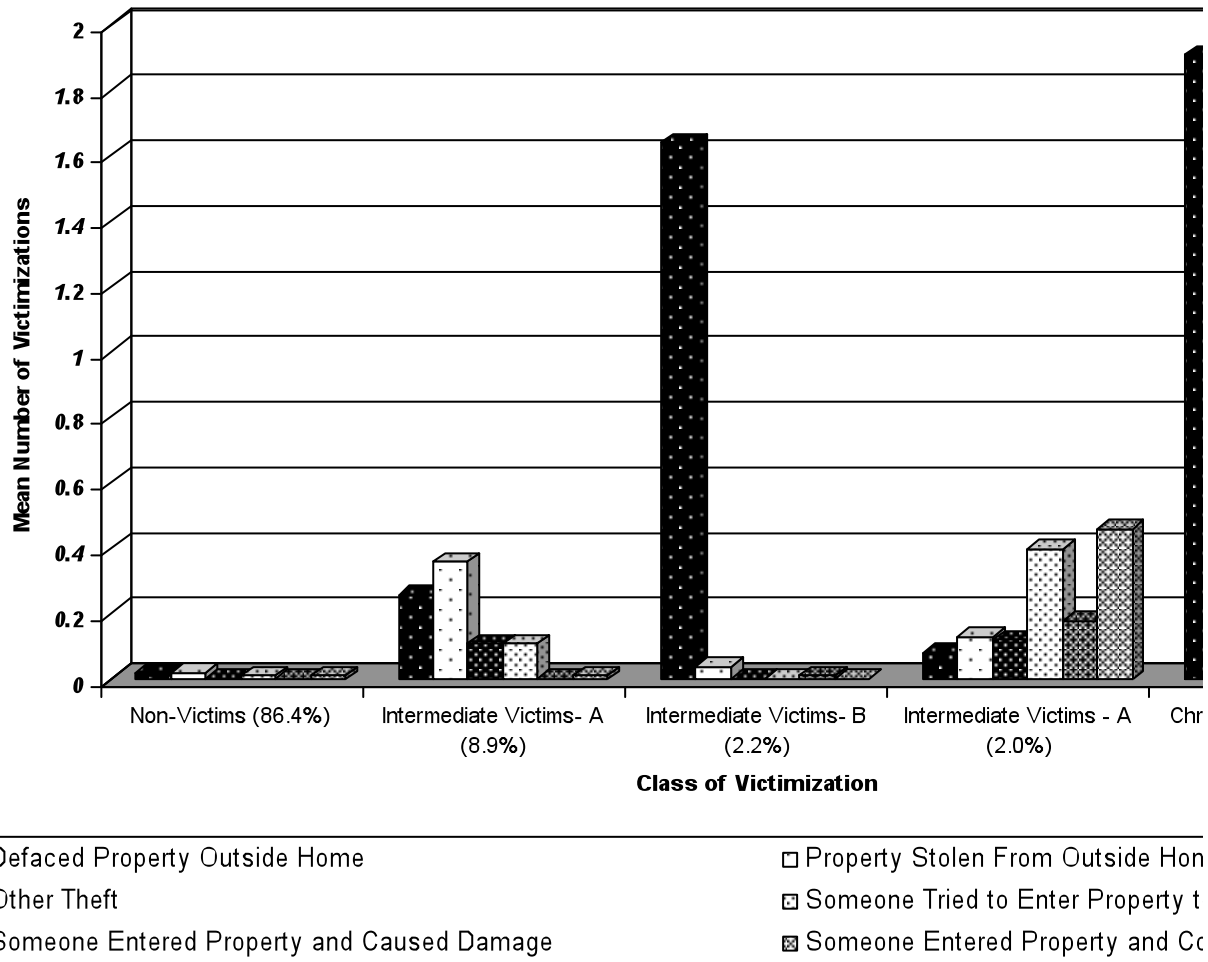


Figure 7: Mean Number of Victimizations by Personal Crime Type - Five Group BCS Model

Heterogeneity in the Frequency Distribution of Crime Victimization

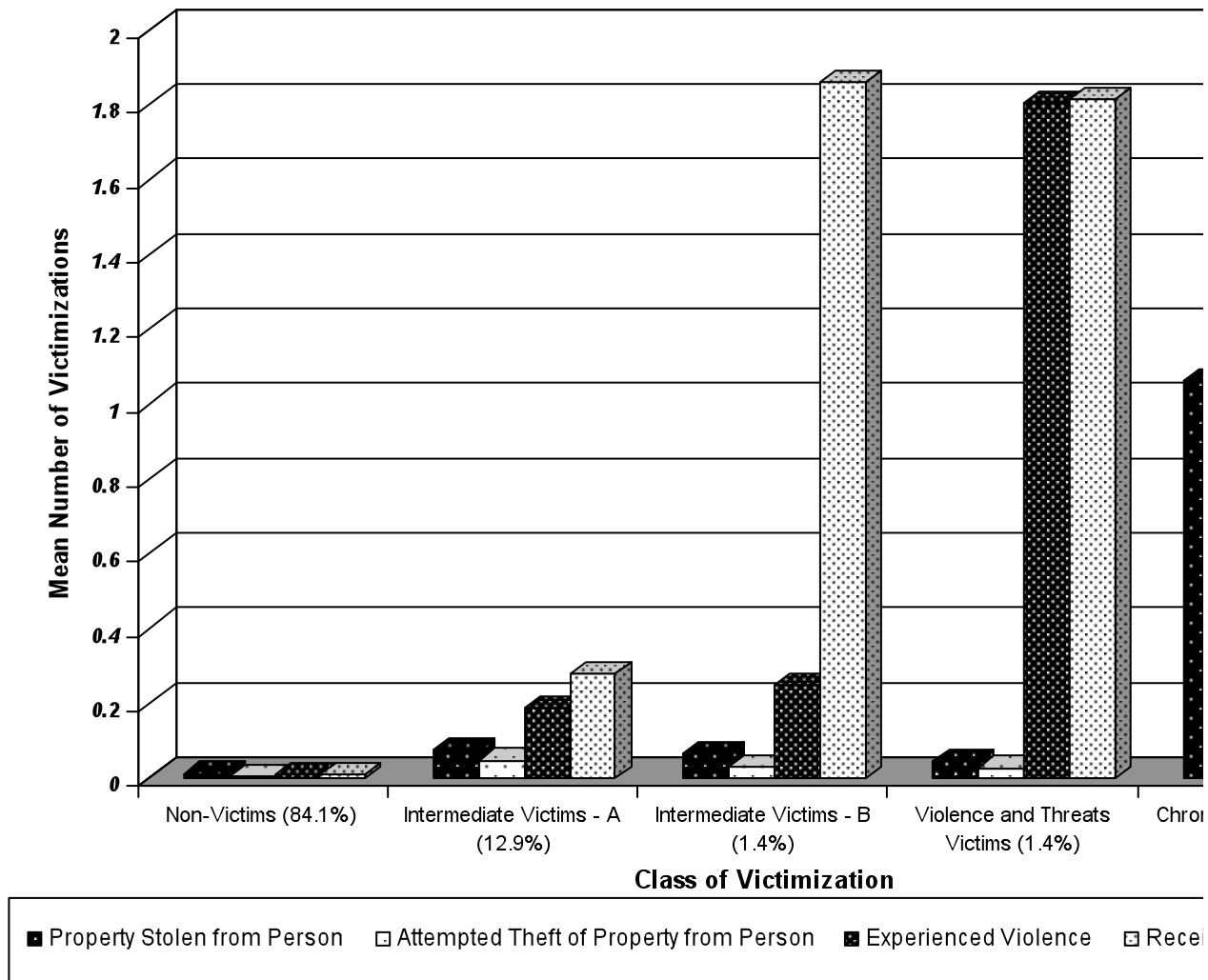


Figure 8: Mean Number of Victimization by Personal Crime Type - Four Group SCVS Model

Heterogeneity in the Frequency Distribution of Crime Victimization

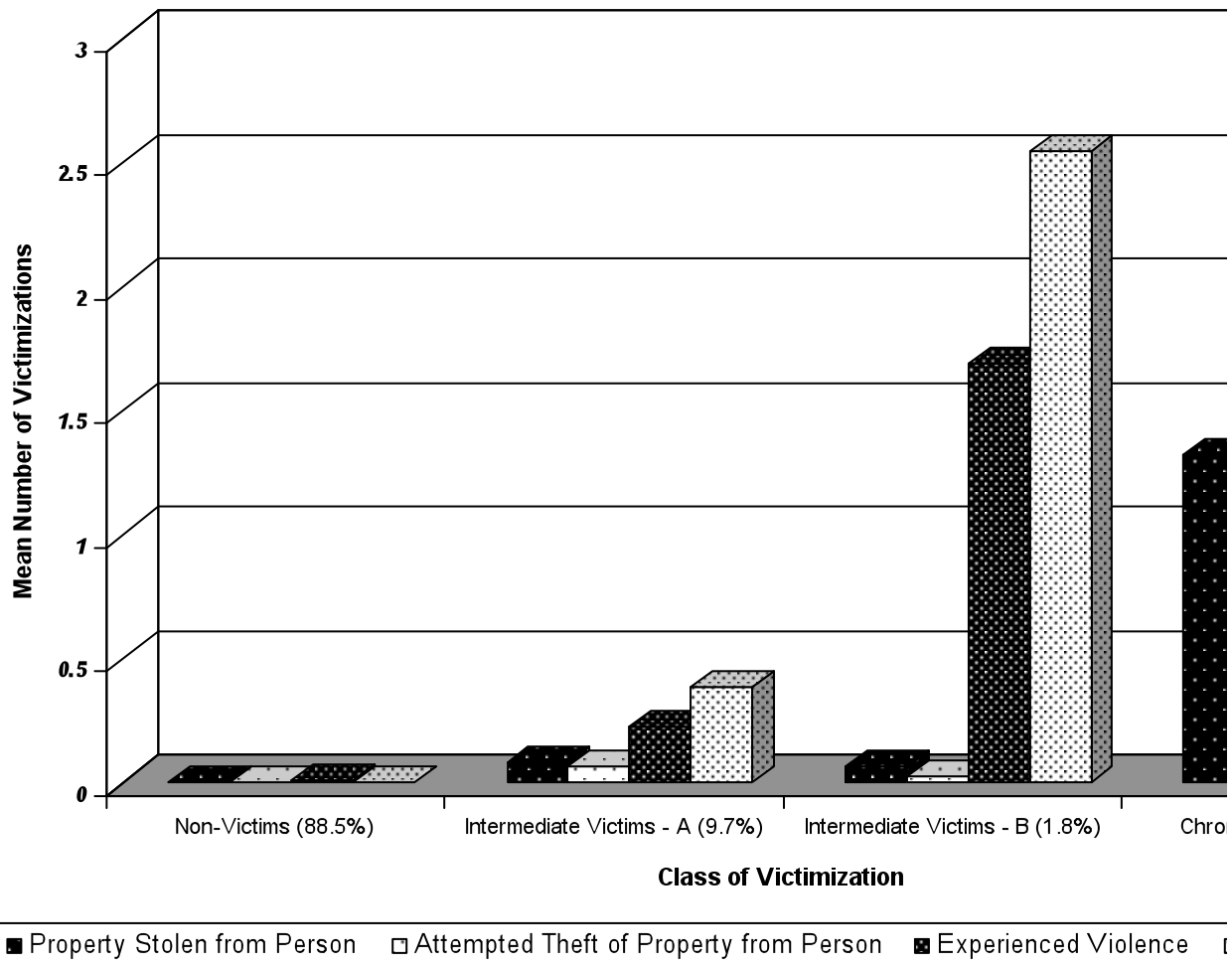


Figure 9: Cumulative proportions of population and victimization by class (BCS Property Six Class Solution)

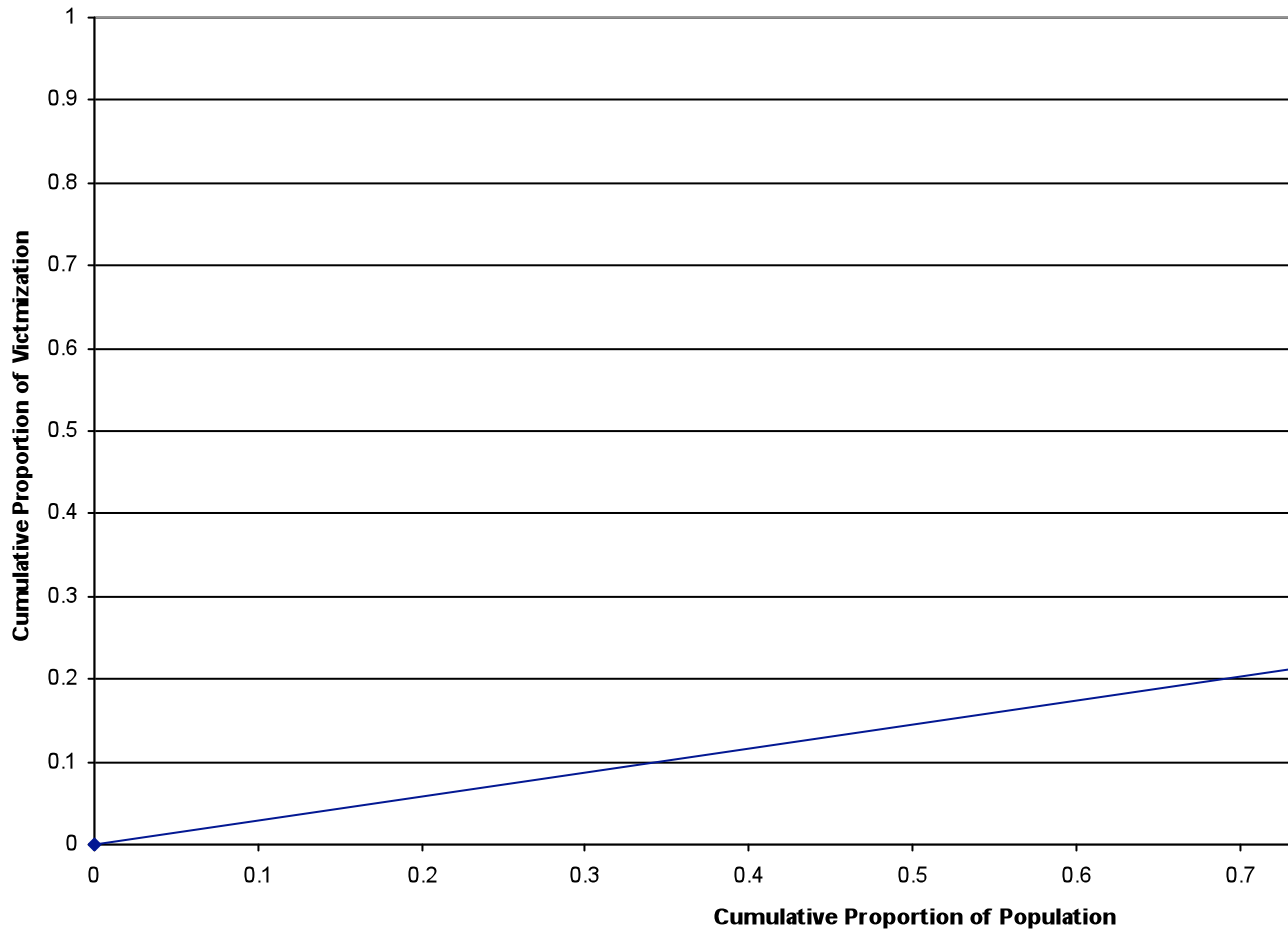
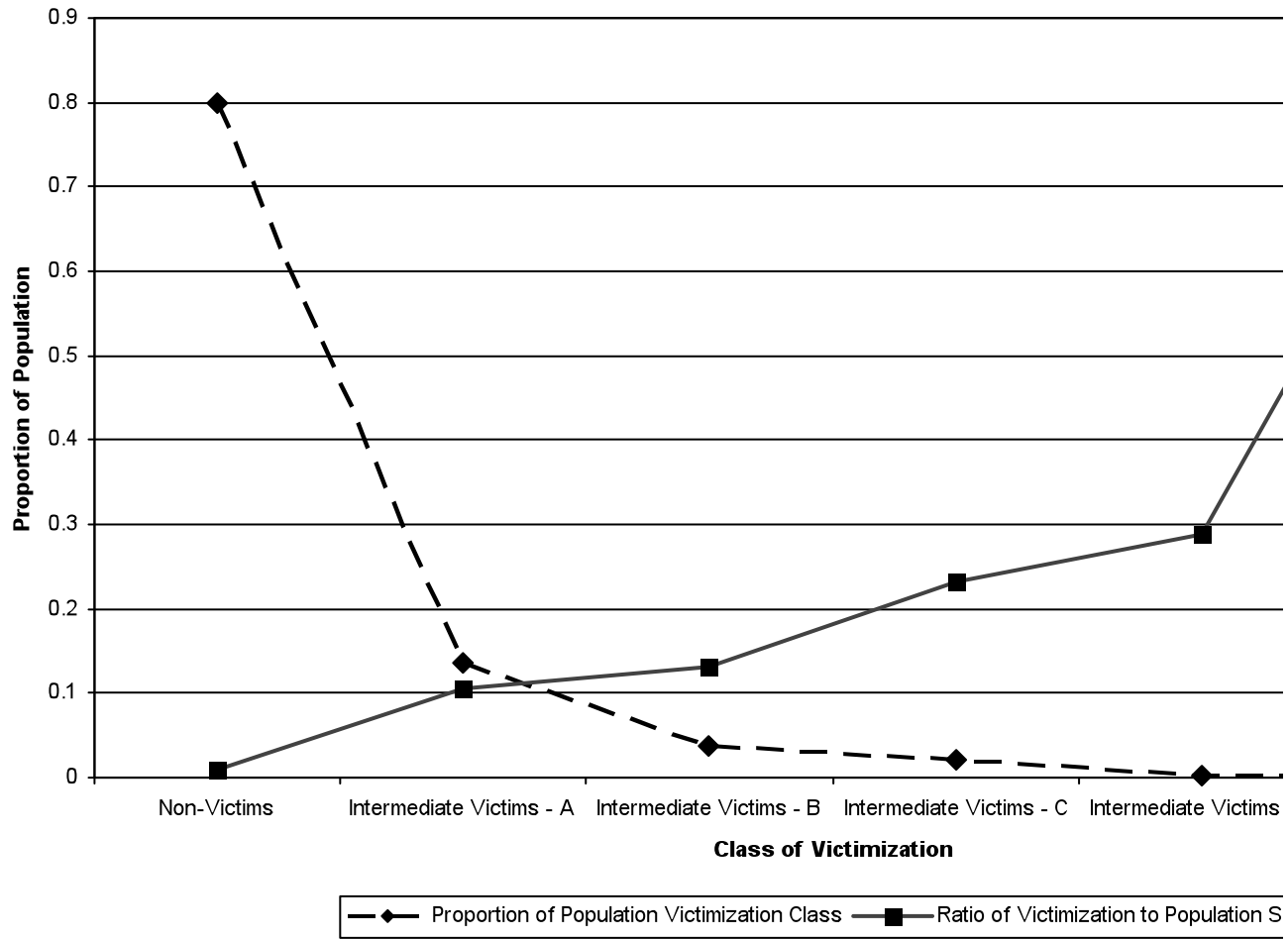


Figure 10: proportion of population and mean victimization by class (BCS Property Six Class Solution)

Heterogeneity in the Frequency Distribution of Crime Victimization



Heterogeneity in the Frequency Distribution of Crime Victimization

Table 2: Cross-tabulation of Membership in Five Class and Six Class Models of BCS Property Crime (percentages)

		Six Class Model					
		Non-Victims	Intermediate Victims - A	Intermediate Victims - B	Intermediate Victims - C	Intermediate Victims - D	Chronic Victims
Five Class Model	Non-Victims	99.7	0.0	0.0	0.0	0.0	0.0
	Intermediate Victims - A	0.0	42.0	1.3	0.0	0.0	0.0
	Intermediate Victims - B	0.3	0.0	84.1	1.2	72.0	16.6
	Intermediate Victims - C	0.0	58.0	12.2	98.2	11.9	1.0
	Chronic Victims	0.0	0.0	2.4	0.6	16.1	82.4
	All Victims	100	100	100	100	100	100

Figure 11: Probability of Membership in the 'Polar Classes' Conditional on Frequency of Victimization. BCS Property crime

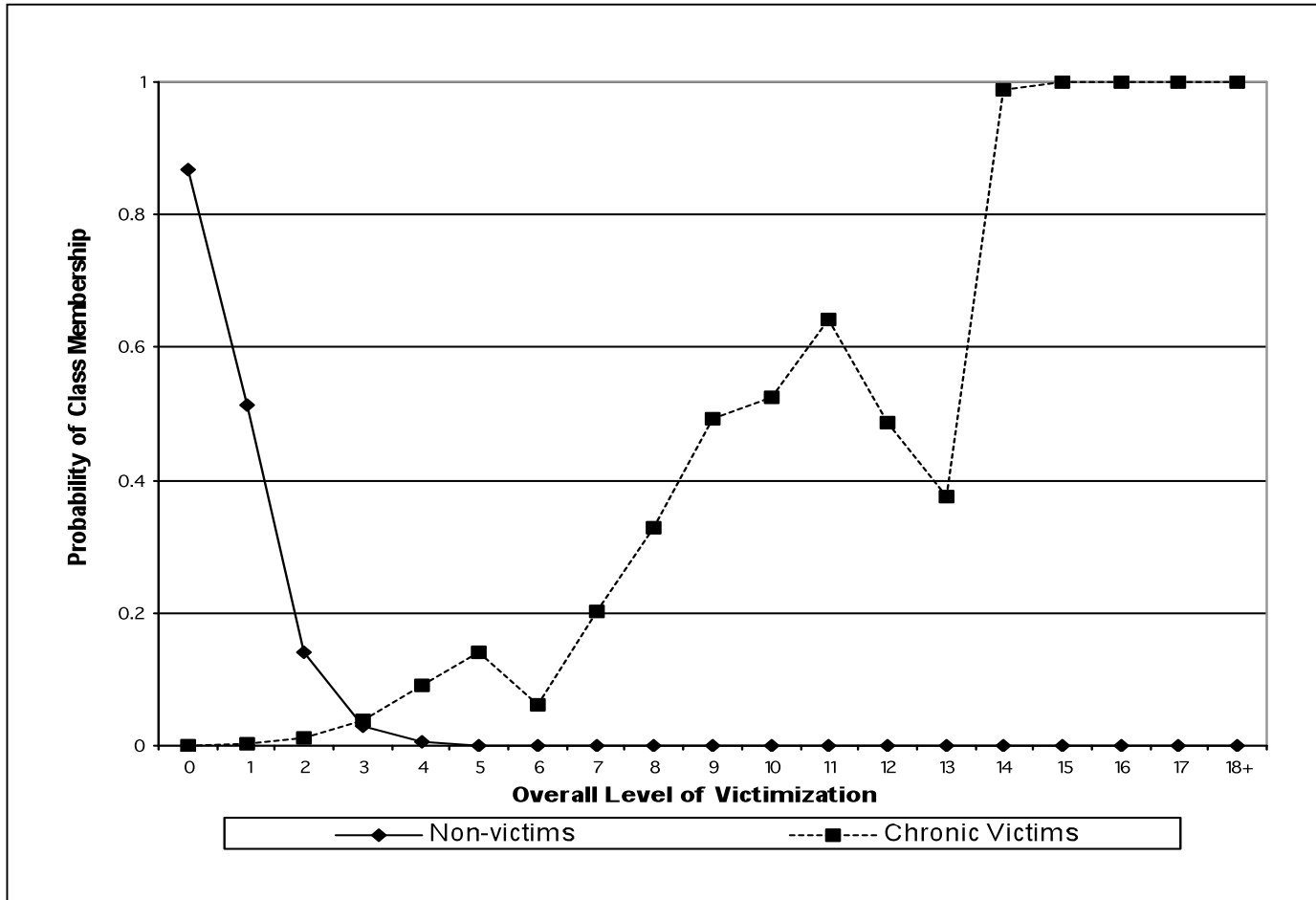
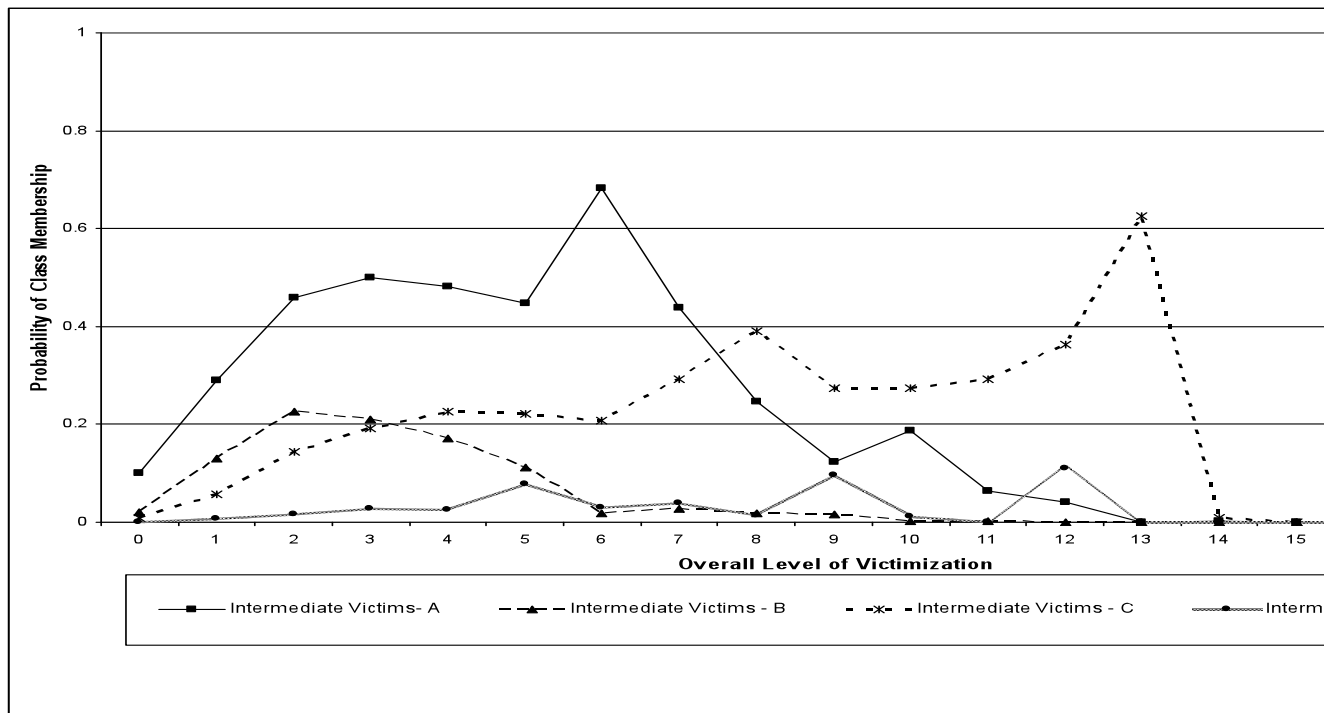


Figure 12: Probability of Membership in the 'Intermediate Classes' Conditional on Level of Victimization. BCS Property Crime



Heterogeneity in the Frequency Distribution of Crime Victimization

Table 3: Class Membership as a Function of Number of Capped Victimization Counts – BCS Property 6 Class Solution

Number of Capped Counts	Number of respondents	Non-Victim Membership Probability	Intermediate Victims - A Membership Probability	Intermediate Victims - B Membership Probability	Intermediate Victims - C Membership Probability	Intermediate Victims - D Membership Probability	Chronic Membership Probability
0	121,554	0.800	0.135	0.038	0.021	0.002	0.002
1	380	<0.001	0.644	0.011	0.228	0.038	0.077
2	9	0.000	0.049	0.001	0.373	0.130	0.446
3	2	0.000	0.000	0.000	0.000	0.000	1.000
4	1	0.000	0.000	0.000	0.000	0.000	1.000
Uncapped mean per person		0.07	4.7	2.28	5.37	6.52	11.57
Capped mean per person		0.07	0.73	0.92	1.62	2.02	5.59
Ratio of uncapped to capped means		0	6.44	2.48	3.31	3.28	2.07

Data are unweighted. N = 121,946. For respondents who said there were “too many” victimizations to recall, their counts were set to missing and then recoded to the mean of those respondents reporting a frequency count of 7 or higher.

Figure 13: Probability of class membership and distribution of property crime victimization: BCS Two-class solution

