Impact of Influential Observations on Enumeration and Variance Estimation in the National Crime Victimization Survey

Michael Planty
Lynn Langton

Bureau of Justice Statistics
810 7th St., NW Washington, DC 20531
Michael.Planty@usdoj.gov
Lynn.Langton@usdoj.gov

Abstract

The National Crime Victimization Survey (NCVS) collects information from respondents about the number and type of victimizations they experience each year. A small percentage of respondents experience a high number of victimizations within the reference period, referred to as series incidents. These chronic victims are substantively important groups as they represent victims of violence from intimate partners and incidents that occur routinely in the workplace and at school. These victims, however, introduce the possibility of extreme weights in the NCVS data. These extreme weight outliers can be influential and significantly increase both point estimates and variance of estimates. One way to control for these impacts is to trim and smooth extreme weights. Trimming weights has the potential to increase bias, but the increase in bias can be offset by the decrease in variance. This paper describes various approaches to incorporating series incidents into the enumeration scheme through trimming while balancing concerns about point estimates and variance.

1. Introduction: Outliers and influential or extreme observations

Outliers in samples are defined as data points that deviate substantially from other observations in a variable or population (Chamber 1986; Dixon 1950; Eltinge 2006). Three general types of outliers include: gross measurement errors which are observations that are not true values; representative outliers which are true values and not entirely unique to the population; and non-representative outliers which are considered unique to the population (Chambers, 1985; Eltinge, 2006; Dixon 1950). Similarly, Anscombe (1960) identified outliers arising from errors in the data and those arising from the inherent variability of the data. The sources of outliers include data errors (i.e., coding errors), respondent response errors (i.e., intentional errors or misreporting due to misunderstanding the question or response categories), sampling error (e.g., sampling units outside of the target population or scope), interviewing or field collection operations, and legitimate cases from the correct population (a true value in the tail of a distribution). Other variations include inliers and fringeliers which tend to be more difficult to identify through simple univariate distributions (e.g., Wainer 1976; Chamber 1986; Eltinge 2006). These outlier types do not appear as outliers when full population estimates are produced, but emerge as outliers for subdomain estimates.

Outliers are concerning because these observations can potentially distort point estimates and inflate variance estimates, increasing volatility and reducing the statistical power of the data. Outliers with extreme weights can disproportionately influence the overall estimate or subdomain estimates. This results in the potential for volatile shifts from year to year and instability in estimates caused by an increase in variance due to the processing and weighting procedure.

However, putting aside gross errors in the collection or editing process, discarding outlier observations is not a straight-forward process. There is a need to balance legitimate true values and the reduction of bias with a reduction in variance. To address such concerns, the statistical field has focused on the identification and handling of extreme influential observations. As Eltinge (2006) noted, to be considered an outlier the observation must deviate from the norm but it also must be influential by having "a major effect on analysis of the data." Techniques for identifying extreme influential observations include a reliance on technical standards related the standard deviations from the

mean, visual inspection, and more sophisticated statistical approaches to identifying influential cases. Solutions to addressing outliers due to coding and measurement errors focus on design and field operations to improve response and processing accuracy (e.g., improving item and questionnaire designs; validating responses in the field; implementing logic checks; eliminating hand coding or type errors through automation). Approaches to handling outliers in the processing (e.g., weighting) and analysis phases including elimination, transformation, capping, trimming, and imputation (e.g., Little; Barnett and Lewis, 1994).

This paper focuses on one set of extreme observation from the National Crime Victimization Survey (NCVS) that impacts the counting of criminal victimization, victims who suffer high-volume, repeated victimizations or series incidents. Series incidents occur when a victim experiences a large number of similar crimes within a given reference period. Research has demonstrated the validity of these observations as true values and has found them to be substantively important to the level and characterization of crime (Lynch, Berbaum, and Planty, 1998; Planty and Strom, 2007; Dodge, 1984a; 1984b; 1987; Lauritsen et al. 2012). NCVS data collection procedures reduce the collection burden associated with series incidents, but these procedures can result in processing, enumeration, and classification issues. In addition, decisions related to how the final weighting procedure will handle series incidents can result in extreme weights.

The paper provides an overview of the series incident collection procedure and the need to account for these incidents in national estimates of crime, and then considers various weighting options for handling series incidents for enumeration and statistical testing purposes. The weighting options are compared and assessed in terms of their impact on violent crime counts, estimate variance, and the coefficients of variation. While there are legitimate concerns about the impact these relatively rare series incidents have on point estimates and variance, eliminating these accurate accounts of crime may also result in distortions in the true crime rate and the characterization of crime victims and incidents.

2. Victimization estimates and series incidents

The measurement and enumeration of the number and type of criminal victimizations occurring during a year is complicated by the challenges associated with repeat victimization. Many crimes types are discrete events with a defined beginning and end. For example, an offender in an armed street robbery may take only a few seconds to threaten, harm, and steal property from a victim. The incident is clearly defined in space and time. In addition, most persons are victimized only once in a given year, if at all. These single, discrete events are relatively easy to recall and describe.

However, for some victims, crime may be a continuous condition (Biderman, 1980; Skogan, 1981). For example, a woman may suffer periodic violence within an on-going intimate partner relationship or victims of bullying at school may suffer continuous assaults during the school year. Victims in these situations may have difficulty defining and recalling the details of each victimization (Dodge, 1984), and these cognitive challenges become apparent when respondents are asked to recall their victimizations during victim surveys. Yet details about each incident are necessary for counting and classifying victimizations that occur each year. Another challenge with these types of high-frequency repeat victimizations, also known as series incidents, is that they can have a significant impact on point estimates and the potential for measuring annual change. A small number of repeat victims can be very influential on the size and precision of the national estimates of crime.

Most crime surveys have data collection limits or a cap on the number of incidents recorded during an interview or in the estimation process in an effort to balance measurement error and accuracy with burden and cost (Lauritsen et al. 2012). The British Crime Survey (BCS), for example, allows up to five incident reports to be completed for any respondent. Respondents who reported being victimized more than once are subsequently asked if they would identify the victimization as a series of crimes that were very similar, done under the same circumstances, and probably by the same people (Bolling, Grant, and Donovan, 2009). When estimating victimization rates, BCS publications typically limit the inclusion of series victimizations to a maximum of five. The national survey conducted for Mexico also uses a cap of five incidents when estimating victimization rates (Encuestas Nacionales Sobre Iseguridad), while the Canadian Victimization Survey limits the number of incident reports to 10 per respondent for each individual crime type and to a total of 20 per respondent. In addition, the Canadian survey also utilizes a series protocol, incorporated in the survey's weighting procedure, which caps the number of similar incidents in the series at three (Nazaretian, Z., & Merolla, D. M., 2013).

The U.S. NCVS utilizes an alternative protocol to cap series incidents. If the respondent was victimized six or more times during the reference period in a similar manner (i.e. by the same offender or in the same general sort of way) and cannot recall the details of every incident, the interviewer uses the series victimization protocol, recording the number of victimizations that occurred and collecting detailed information for the last occurrence. Since specific information used to classify the crime type is only gathered for the most recent victimization, all incidents in the series are classified by the most recent incident. The series protocol addresses the recall and burden issues for the respondent, but at the cost of having less available information about these experiences in the data. It can also result in a highly skewed distribution of the number of incidents experienced by victims (figure 1, 2).

These series incidents are not considered errors, but rather true estimates of victimization with substantive importance (Planty and Strom, 2007; Lauritsen et al. 2012). When examined by crime type and incident characteristics we find there are substantively important instances of repeat victimization. Prior research examined the characteristics of violent series victimizations and found series consisted of three typical categories of assaults: those that occurred in the home and involved intimate partners or domestic violence; those that occurred at school; and those associated with workplace violence (Lynch, Berbaum, and Planty, 1998; Planty and Strom, 2007; Dodge, 1984a; 1984b; 1987). Further, series incidents are not periodic anomalous outliers occurring occasionally, but rather, are consistently reported each year and under similar conditions. The fact that the NCVS procedures routinely capture these victims is evidence of a real phenomenon and not a gross measurement error (figure 3, 4).

2.1 NCVS weighting and capping series incidents

Despite evidence that high-frequency repeat victimizations are a real and important phenomenon, inclusion of these series incidents can have a disproportionately high impact on estimates of criminal victimization. Capping these incidents and limiting the number of separate criminal incidents that can be included in the total estimates of crime clearly reduces the resulting survey estimates. Therefore, efforts to place an upper limit on the number of incident reports that can be collected in the field or to restrict the inclusion of series incidents during data processing or analysis have been criticized by some researchers for undercounting the extent of crime victimization (Planty and Strom, 2007; Farrell and Pease 2007). However, capping remains a common practice applied in most victimization surveys and is regarded by survey methodologists as a means to provide more consistent comparisons and trend measures that are less impacted by relatively rare extreme outliers. The question then becomes how to best cap the estimate in a manner that balances estimate accuracy and volatility.

At various points in the past decade, the NCVS has utilized several different approaches to handling series incidents. One practice was to exclude series incidents completely from all estimates of crime in the NCVS annual Criminal Victimization bulletin. The exclusion approach was applied because of concerns that classifying all incidents in the series according to the most recent could result in some misclassification. In other NCVS analyses the entire series was counted as one incident, to ensure that the single, most recent incident was correctly classified and that characteristics of that incident were captured in estimates. Finally, acknowledging that these prior approaches underestimated the annual crime numbers, the final weight has also been adjusted to count the actual number of incidents in in the series, up to a maximum of ten.

3. Methodology

This paper examines several capping strategies for the NCVS series incidents: 1. Exclude entirely (count as 0); 2. Count as one incident; 3. Cap the estimate at the median response for a given year; or 4. Record the actual number of incidents as reported by the respondent. The paper then compares point estimates generated from each of these approaches for the major violent crime counts: total violence; rape or sexual assault; robbery; aggravated assault; and simple assault. In order to measure of estimate volatility, it also examines the coefficient of variation (CV) or relative standard error (RSE) of each estimate, which is the ratio of the standard error to the mean, with larger CVs meaning less precision.

3.1 Analysis

The analysis is based on NCVS data from 1993-2012. For the purpose of examining the handling of series incidents,

the data are restricted to violent crime, which includes attempted and completed rape, sexual assault, robbery, and both aggravated and simple assault. In the NCVS, the type of crime classification is based upon detailed characteristics of the event as provided by the respondent. During data processing an algorithm classifies each incident into a type of crime, depending on the victim's responses to a number of items on the survey questionnaire.

The NCVS sample estimates are weighted to represent the national person (12 or older) and household populations in a given year. While each component of the weighting process could result in an outlying and influential observation, the analysis primarily focuses on the last stage, series incident adjustment. The weighting steps include:

Components of the NCVS Weights	Household-Level Estimates			Person-Level Estimates		
	Household	Victimization	Incident	Person	Victimization	Incident
Base Weight	×	×	×	×	×	×
Special Weighting Adjustments	×	×	×	×	×	×
Household Non-interview Factor	×	×	×	×	×	×
Within- Household Non-interview Factor				×	×	×
First-Stage Ratio	×	×	×	×	×	×
Second-Stage Ratio	×	×	×	×	×	×
Bounding Adjustment		×	×		×	×
Multiple Victim Adjustment						×
Series Incident Adjustment		×			×	

4. Findings

Before examining the impact of series enumeration rules on counts of violent crime, the first step was to examine the basic descriptive statistics for series incidents and the impact of the various series weighting approaches on these descriptives. The 2,269 victims of high-frequency repeat violent victimizations from 1993 through 2012 experienced an average of 21 incidents and a median of 10 incidents in a six month period (table 1). Estimates of skewness and kurtosis suggest that the data are highly skewed with extreme deviations. The number of incidents in the series ranged from six to 750.

When looking next at the weighted number of victimizations contributed by series incidents for each of the last three weighting options, the impact of these series incidents becomes more apparent (table 2). With series set equal to one the variance is much lower than with series capped at the median or with the actual number of series incidents used. However, the number of victimizations is also substantially lower. As seen in figure 5, the mode and median sample case weight is fairly consistent across the three approaches for handling series cases, while, as expected, the mean is

highest when the actual number of events is used.

When the four weighting options are applied to generate estimates of violent crime counts, the impact of series incidents is even clearer (figures 6, 7, 8). Excluding series incidents or counting them as one incident clearly underestimates the total number of victimizations, though the degree of underestimations was more pronounced in 1993 than in 2012. This is due to the fact that the number of series incidents declined slightly during the 20 year period, as did the number of events that occurred within each series incident (figures 3, 4). While using the actual number of incidents results in the higher crime counts, the CVs are also higher, meaning more variance and greater volatility of estimates. This is especially apparent when examining the rape and sexual assault estimates for a single year (figure 7). The CV when series incidents are capped at ten, the median number of incidents, is under 25%, while the CV is approximately 35% when the actual number is used. However, compared to other crime types, rape/sexual assault estimates have relatively large CVs regardless as to how series are treated.

Figures 9 and 10 and tables 3, 4, 4a, and 4b summarizes the percent change in violent crime estimates and CVs using the different series weighting approaches. Again, estimates of rape and sexual assault are the most sensitive to the inclusion of series incidents. When series incidents are included in counts of rape and sexual assault the counts increase by less than 10% to over 130% depending on the weighting approach used. The increases are less pronounced for crimes like robbery and aggravated assault that have larger sample sizes and are less likely to involve high frequency repeat victimizations. However, even for these crime types, using the capped series weight results in a more than 10% increase in the victimization count and using the actual number of series events results in about a 20% increase.

The need to balance accuracy and increased variance is also apparent from these tables and figures. Across the board the CVs increase when series incidents are capped, rather than set equal to zero or one. When all series events are counted, the CVs are even greater and suggest that for rarer events, like rape and sexual assault, and for crimes were series incidents are more prevalent, the estimates became more volatile and measuring year-to-year change becomes more difficult. However, this problem in detecting annual change for rare events exists regardless of the series treatment. Other approaches, such as generating rolling annual averages, help to stabilize the volatility (Planty et al. 2013).

5. Discussion

The findings presented here suggest that the capping mechanism used to handle NCVS outlier cases contains the volatility of the estimates, results in a one-time level change, and for most estimates results in minor or reasonable increases to the standard errors and measured by the CVs. For example, compared to series counted as one, total violent crime, rape and sexual assault, robbery, aggravated assault, and simple assault 2012 estimates increased between 30 to 60% using the capping approach and saw relatively small increases to the CVs. As expected some estimates do incur substantial increases to point estimates and variance by these influential cases. However, for these cases, the impact of series incidents is driven from estimates already prone to small sample sizes, low statistical power, and high standard errors. For example, the number of rape/sexual assault increases by 59% with an increase of 187% in the standard error. Even without the inclusion of series cases with relatively high standard errors and CVs, this estimate performed poorly, though. The new series capping rule made a poor measure slightly worse.

To summarize:

- High volume repeat victims pose challenges to data collection and estimation
- Reporting actual respondent counts produces extreme and influential weights resulting in:
    - Large increases in point estimates
    - Large shifts in annual change
    - Large increases in variance, reduction in statistical power
    - Relatively rare crimes such as rape/sexual assault experience large increases in volatility, other crime types less susceptible (e.g., robbery)
    - Relatively common incidents such as simple assault account for a large number of series, but less prone to large shifts because of the large number of total observations.
- Capping series incidents with the median estimate reduces the impact on the estimate and variance while balancing substantive accuracy.

This research focused on one particular type of outlier in NCVS data. However, series incidents do not represent the only extreme and influential values in the data. Particularly as the NCVS begins to be used to measure crime at the subnational level, future research is needed to address other issues such as the impact of inliers for subdomains and particular small areas.

Disclaimer

The views expressed here are the author's and do not reflect those of the Bureau of Justice Statistics, U.S. Department of Justice.

References

Anscombe (1960). Rejection of Outliers. Technometrics, 2, 123.

Barnett, V. and Lewis, T.: 1994, *Outliers in Statistical Data*. John Wiley & Sons, 3rd edition.

Chambers, R.L. (1986). Outlier Robust Finite Population Estimation. *Journal of the American Statistical Association* 81, 1063-1069.

Sadeq Chowdhury, Meena Khare, Kirk Wolter (2007). Weight Trimming in the National Immunization Survey Proceedings of the Section on Survey Research Methods, American Statistical Association, http://www.amstat.org/sections/srms/proceedings/y2007/Files/JSM2007-000077.pdf

Dodge, R.W. (1984a). A preliminary inquiry into series victimization. In R.G. Lehnen & W. Skogan (Eds.), *The National Crime Survey working papers, 2.* Washington, D.C.: Bureau of Justice Statistics.

Dodge, R.W. (1984b). Series victimization – what is to be done? In R.G. Lehnen & W.Skogan (Eds.), *The National Crime Survey working papers, 2.* Washington, D.C.: Bureau of Justice Statistics.

Dodge, R.W. (1987). Series crimes: report of a field test. Washington, D.C.: Bureau of Justice Statistics.

Dodge, R.W. & Lentzner, H.R. (1980). Patterns of personal series incidents in the National Crime Survey. *Review of public data use, 8*, 27-41.

Elliott, M.R. and R.J.A. Little (2000). Model-Based Alternatives to Trimming Survey Weights. Journal of Official Statistics 16, 191-209.

Eltinge, J.L. and Cantwell, P.J. (2006): Outliers and influential observations in establishment surveys. Federal Economic Statistics Advisory Committee. http://www.bls.gov/bls/fesacp3060906.pdf

EUICS Consortium (2005). Methodology of the European Crime and Safety Survey. Retrieved from http:// www.europeansafetyobservatory.eu/downloads/WP_methodology.pdf.

General Social Survey Cycle 18: Victimization, 2004 (2005). User's guide to the public use microdata file. Ottawa, Ontario: Statistics Canada.

ICESI (2011). Cuadernos del ICESI 9: Analisis de la Septima Encuesta Nacional Sobre Inseguridad ENSI-7/2010.

Lauritsen, Janet L., Jennifer Gatewood Owens, Michael Planty, Michael R. Rand, and Jennifer L. Truman (2012). "Methods for Counting High-Frequency Repeat Victimizations in the National Crime Victimization Survey." Washington DC: U.S. Department of Justice, Bureau of Justice Statistics, *NCJ 237308*.

Little, R.J.A. and P.J. Smith (1987). Editing and Imputation for Quantitative Survey Data. *Journal of the American Statistical Association* 82, 58-68.

Little, R.J.A., Lewitzky, S., Heeringa, S., Lepkowski, J., Kessler, R.C. (1997).Assessment of Weighting

Methodology of the National Comorbidity Survey. *American Journal of Epidemiology*, 146, 439-449.

Lynch, J.P., Berbaum, M.L., & Planty, M. (1998). *Investigating repeated victimization with the NCVS*. Final report for National Institute of Justice Grant 97-IJ-CX-0027. Washington, D.C.: U.S. Department of Justice.

Nazaretian, Z., & Merolla, D. M. (2013). Questioning Canadian Criminal Incidence Rates: A Re-analysis of the 2004 Canadian Victimization Survey. *Canadian Journal of Criminology and Criminal Justice/La Revue canadienne de criminologie et de justice pénale*, *55*(2), 239-261.

Planty, M. & 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.

Planty, M., Langton, L., Krebs, C., Berzofsky, M., & Smiley-McDonald, H. (2013). *Female victims of sexual violence, 1994-2010.* (NCJ 240655) Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics.

Potter, F.J. (1988). Survey of Procedures to Control Extreme Sampling Weights. Proceedings of the Section on Survey Research Methods, American Statistical Association, 453-458.

Potter, F.J. (1993). The Effect of Weight Trimming on Nonlinear Survey Estimates. Proceedings of the Section on Survey Research Methods, American Statistical Association, 758-763.

Figure 1. Number of victimization per victim reported in the 6-month reference period, NCVS 1993-2012

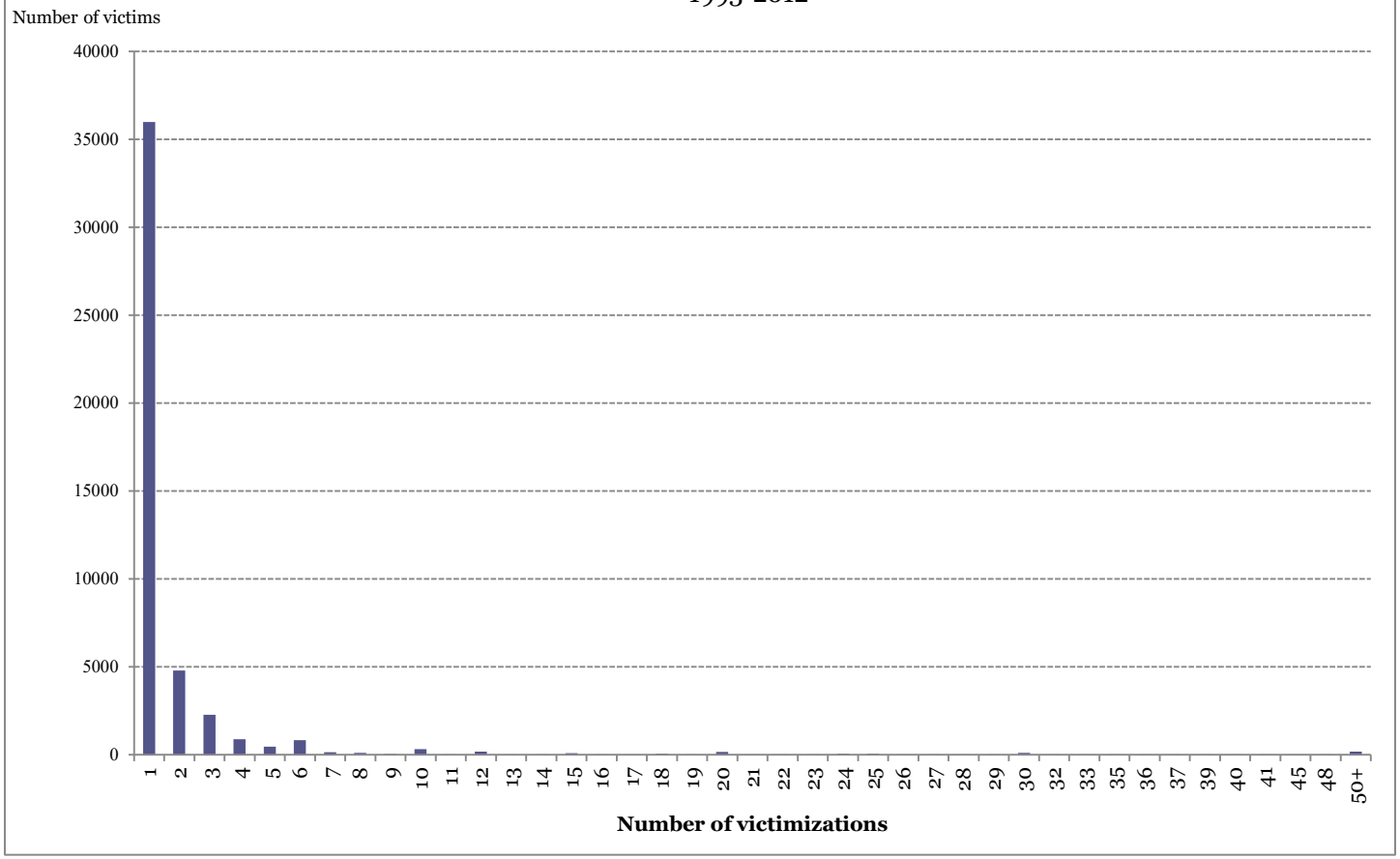


Figure 2. Number of victimizations by series victim, NCVS 1993-2012

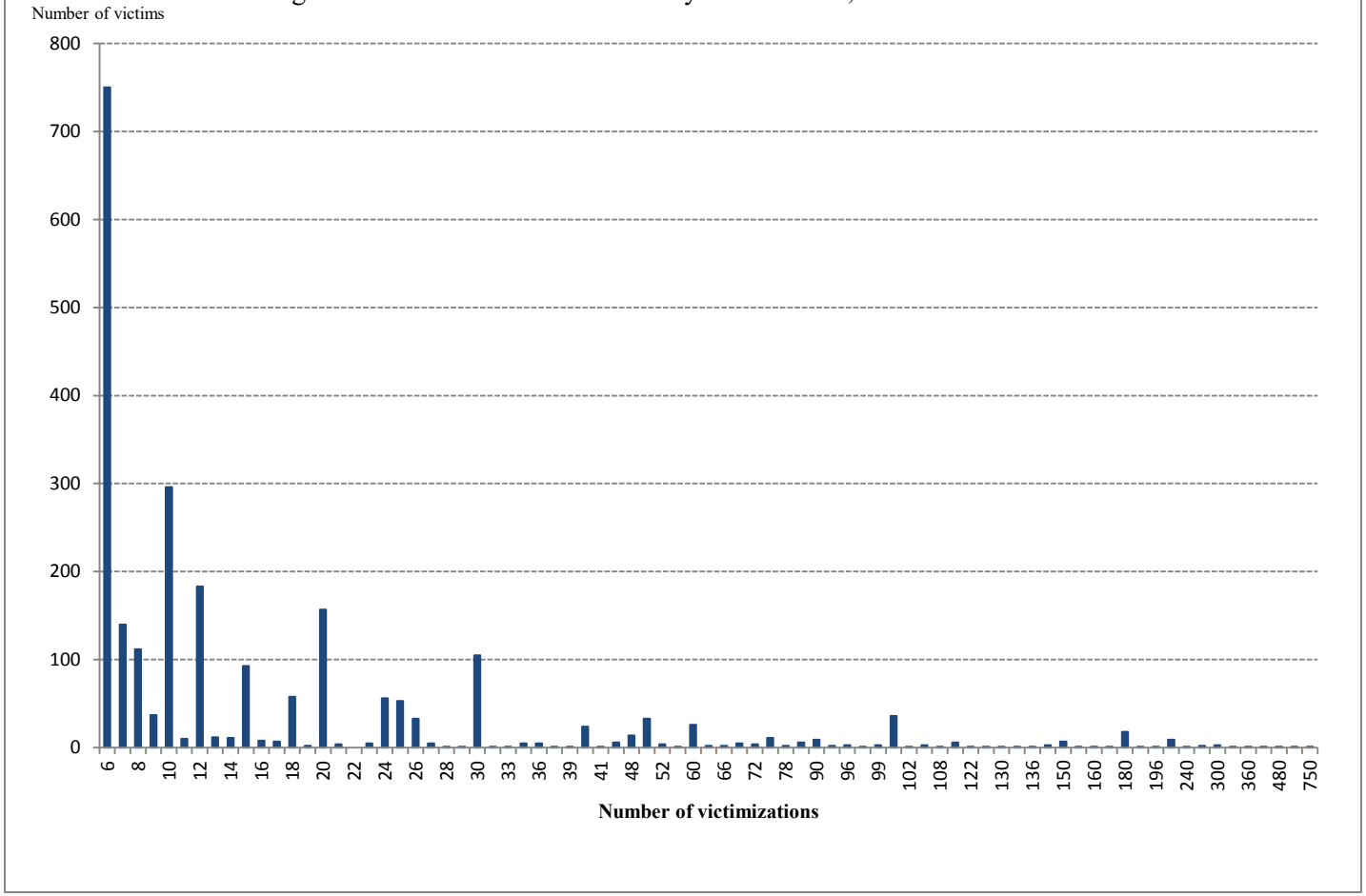


Figure 3. Number of series incidents reported per year, NCVS, 1993-2012

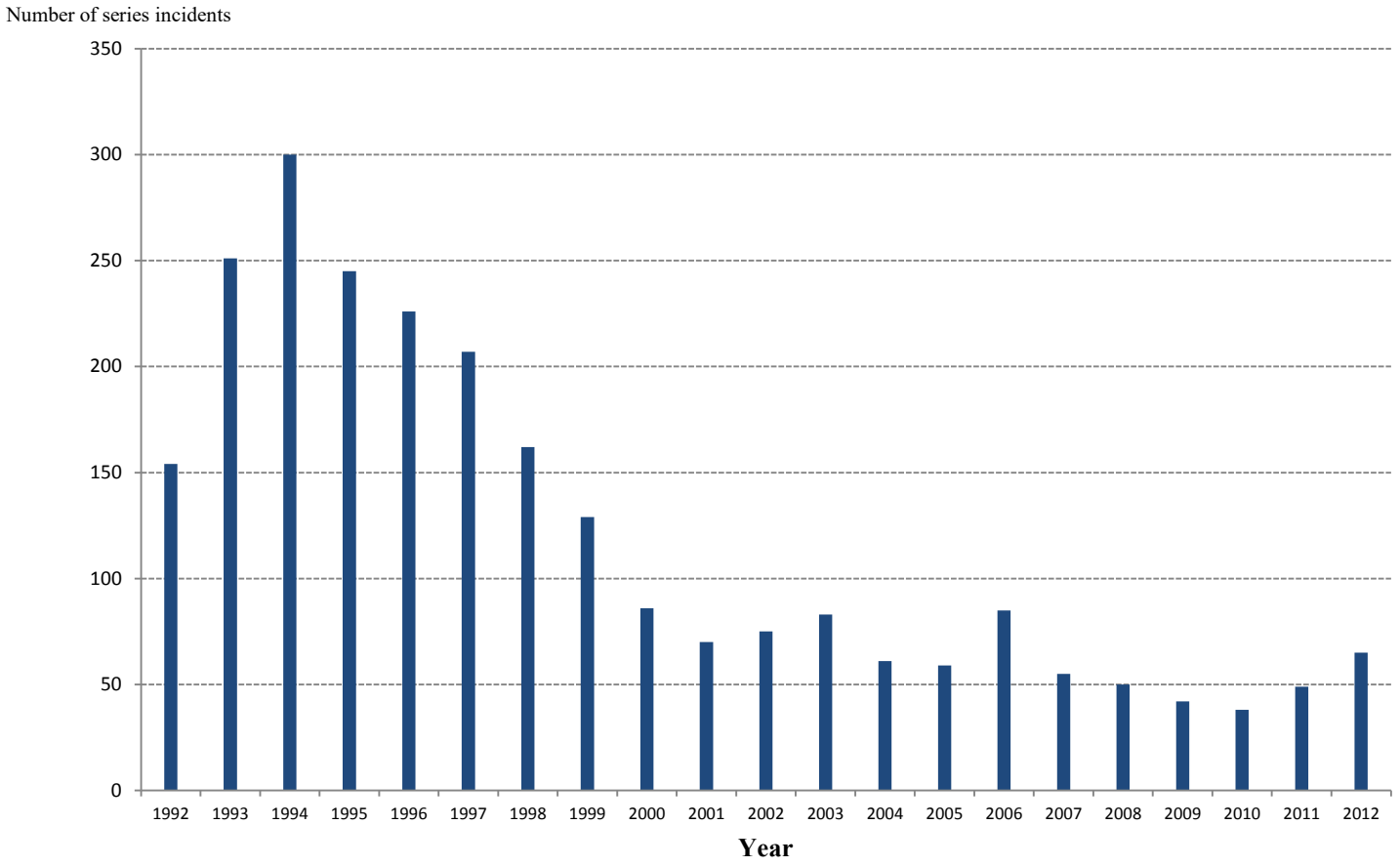


Figure 4. Percent of violent crimes that are series incidents, NCVS, 1993-2012

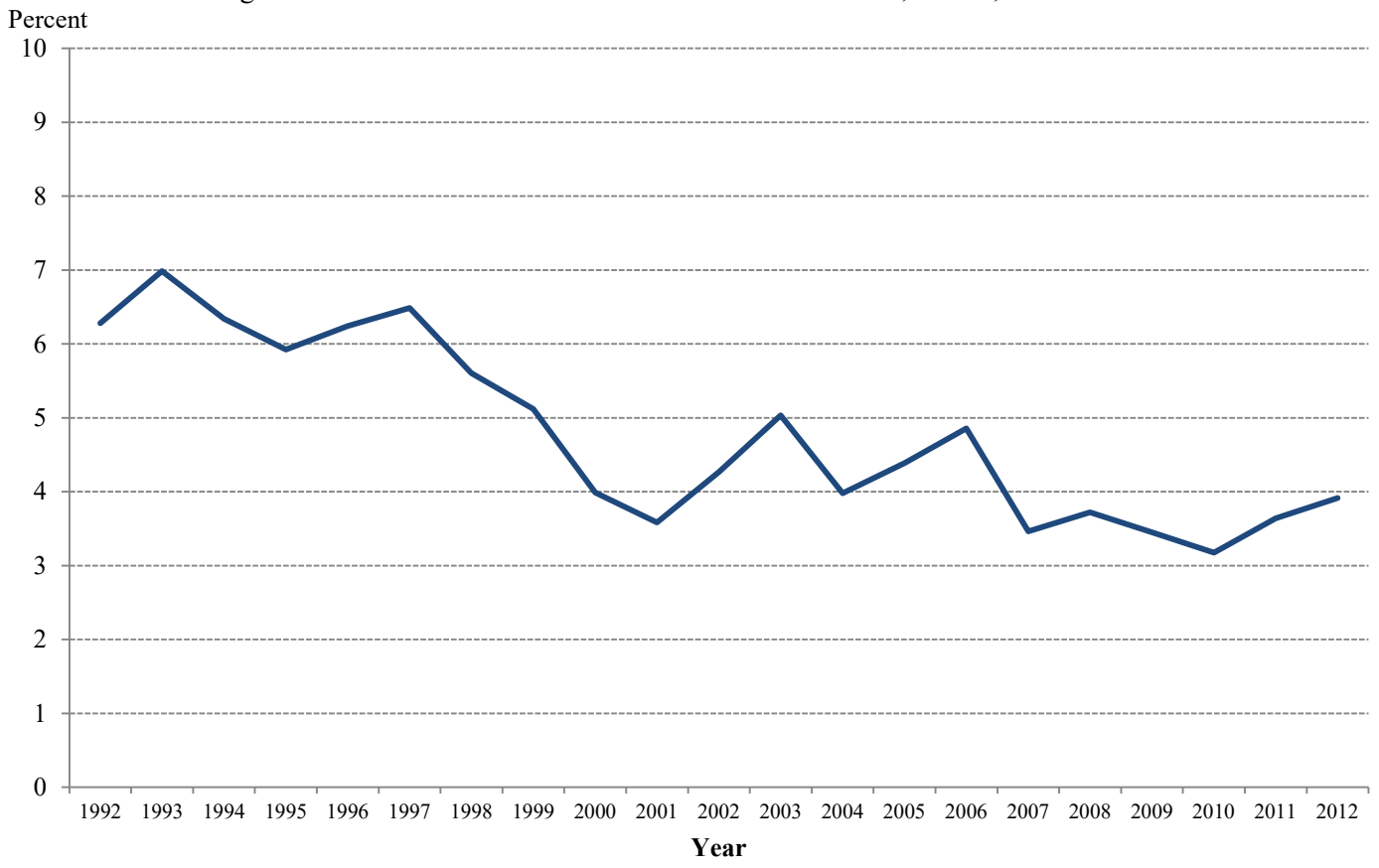


Table 1. Series incident summary statistics, 1993-2012

N		2,269
Mean		21.4
Median		10.0
Mode		6
Std. Deviation		38.9
Skewness		7.4
Kurtosis		88.1
Minimum		6
Maximum		750
Percentiles	25	6
	50	10
	75	20

Table 2. Summary statistics for weighting by series enumeration scheme, 1993-2012

		series=1	series=capped	series=actual
Mean		2,732	3,107	3,621
Median		2,394	2,409	2,408
Mode		2,219	1,977	1,977
Std. Deviation		1,090	3,313	16,924
Variance		1,187,558	10,976,526	286,406,594
Skewness		4	10	71
Kurtosis		36	148	8,244
Range		37,871	153,385	2,733,224
Minimum		101	101	101
Maximum		37,972	153,487	2,733,326
Percentiles	**25**	2,084	2,088	2,088
	50	2,394	2,409	2,408
	75	3,044	3,095	3,094

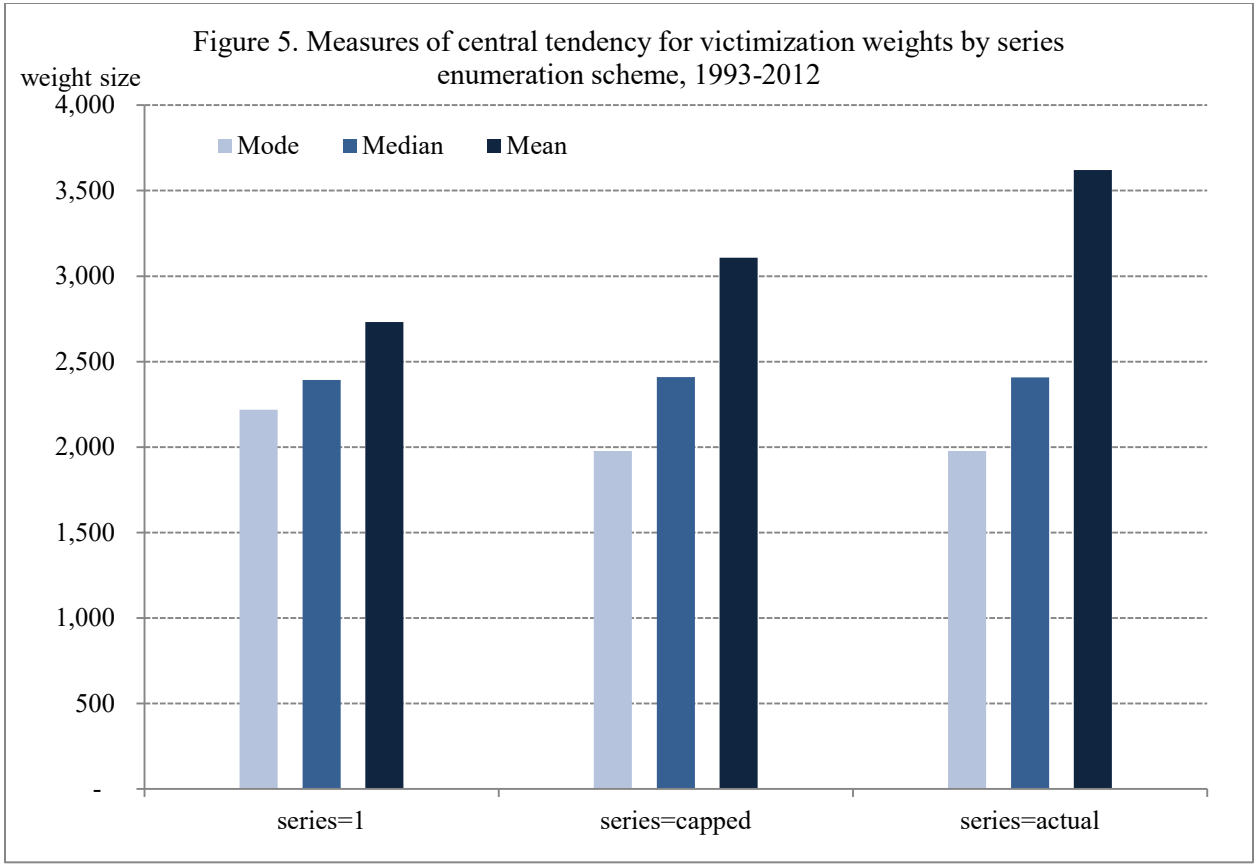


Figure 5. Measures of central tendency for victimization weights by series enumeration scheme, 1993-2012
weight size
4,000
3,500
3,000
2,500
2,000
1,500
1,000
500
-
Mode
Median
Mean
series=1
series=capped
series=actual

Figure 6. Total violence, point estimates and coefficient of variation, 1993, 2012

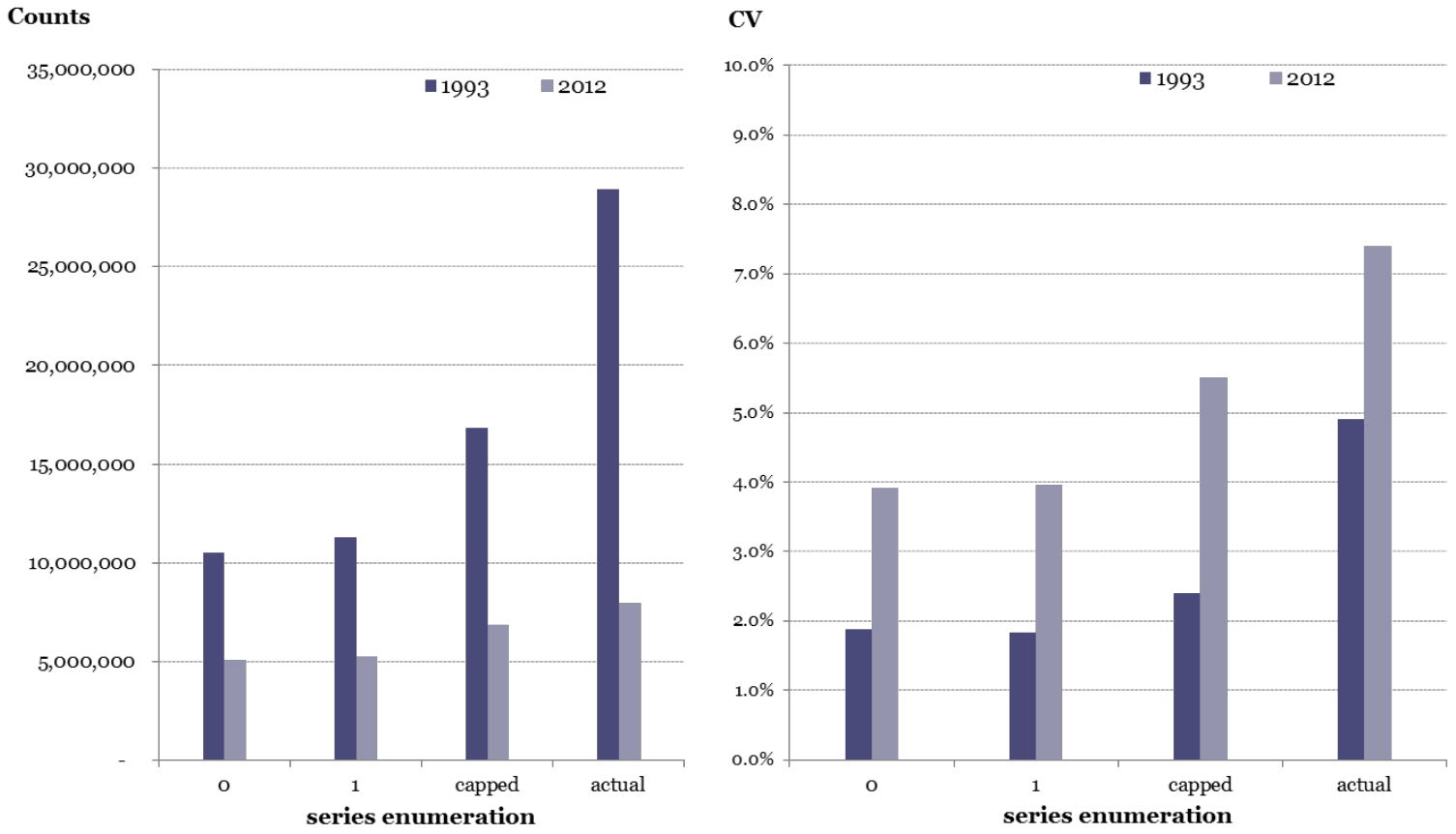
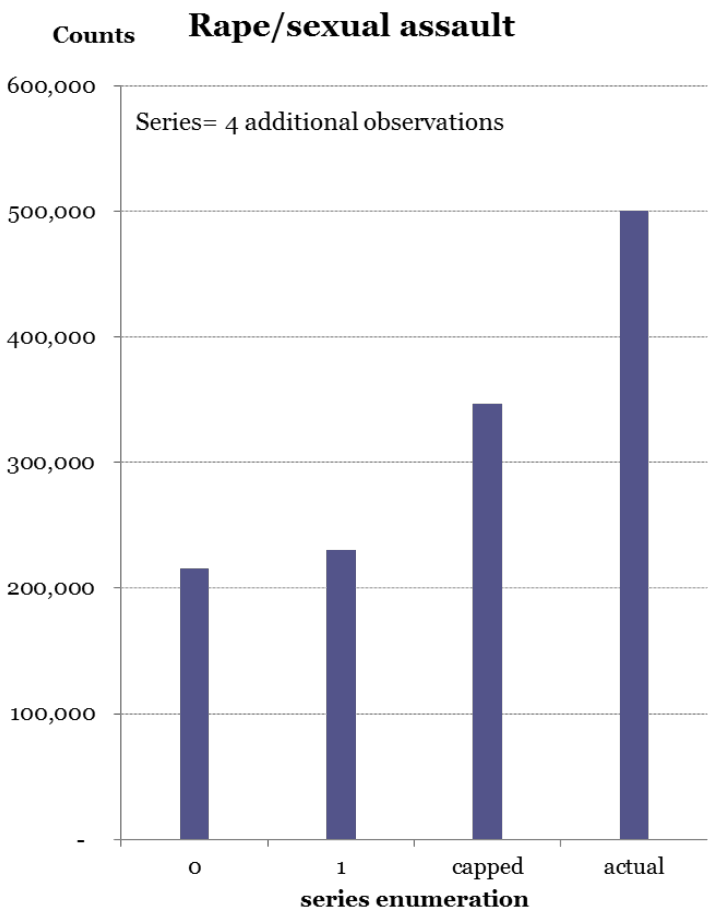


Figure 7. Total rape/sexual assault, point estimates and coefficient of variation, 2012



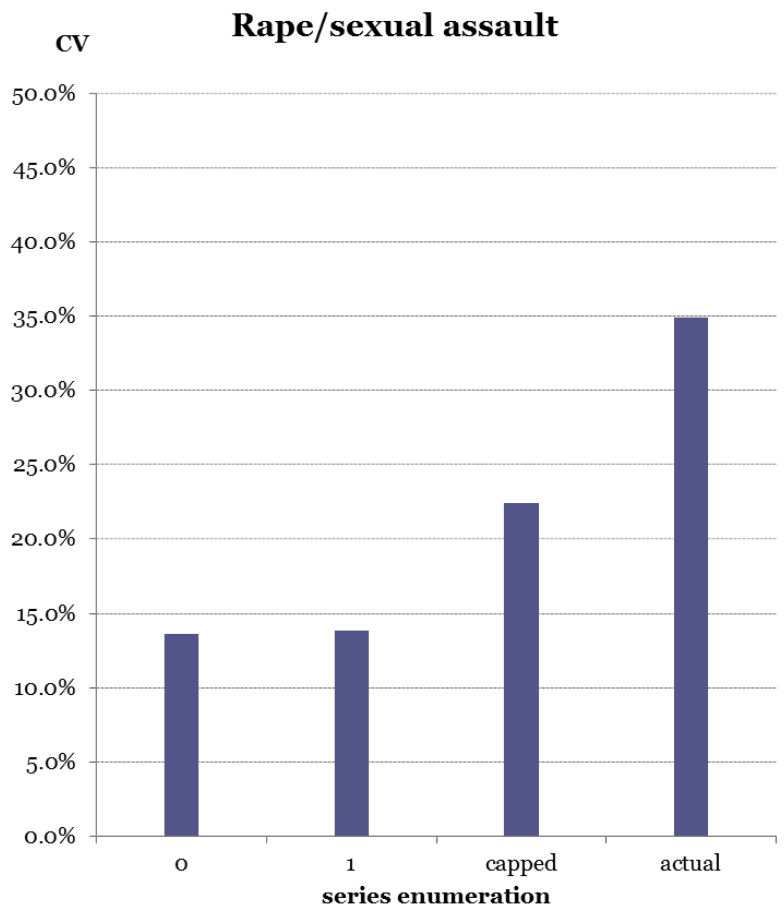


Figure 8. Total simple assault, point estimates and coefficient of variation, 2012

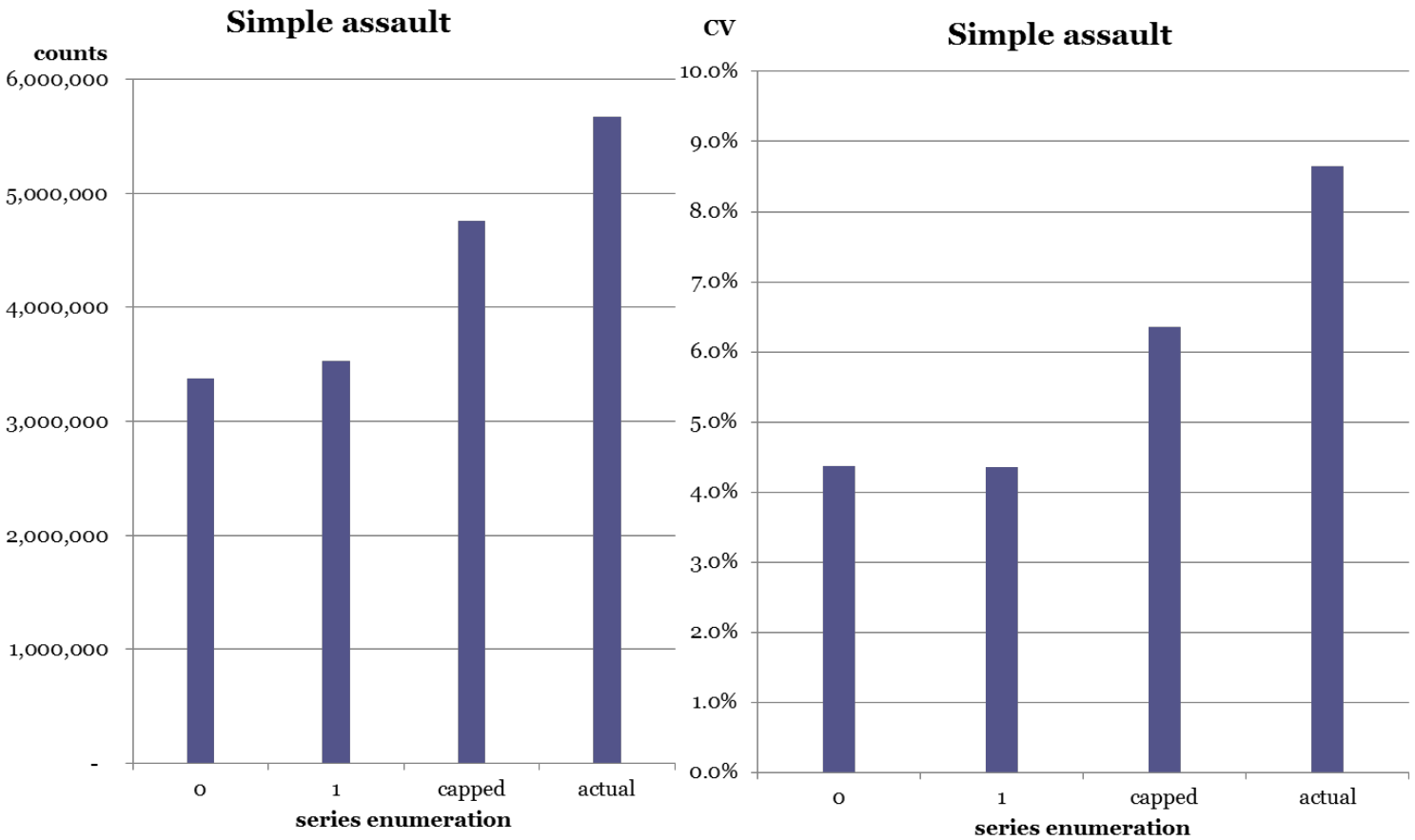


Figure 9. Type of crime, percent increase point estimates (victimization counts), 2012

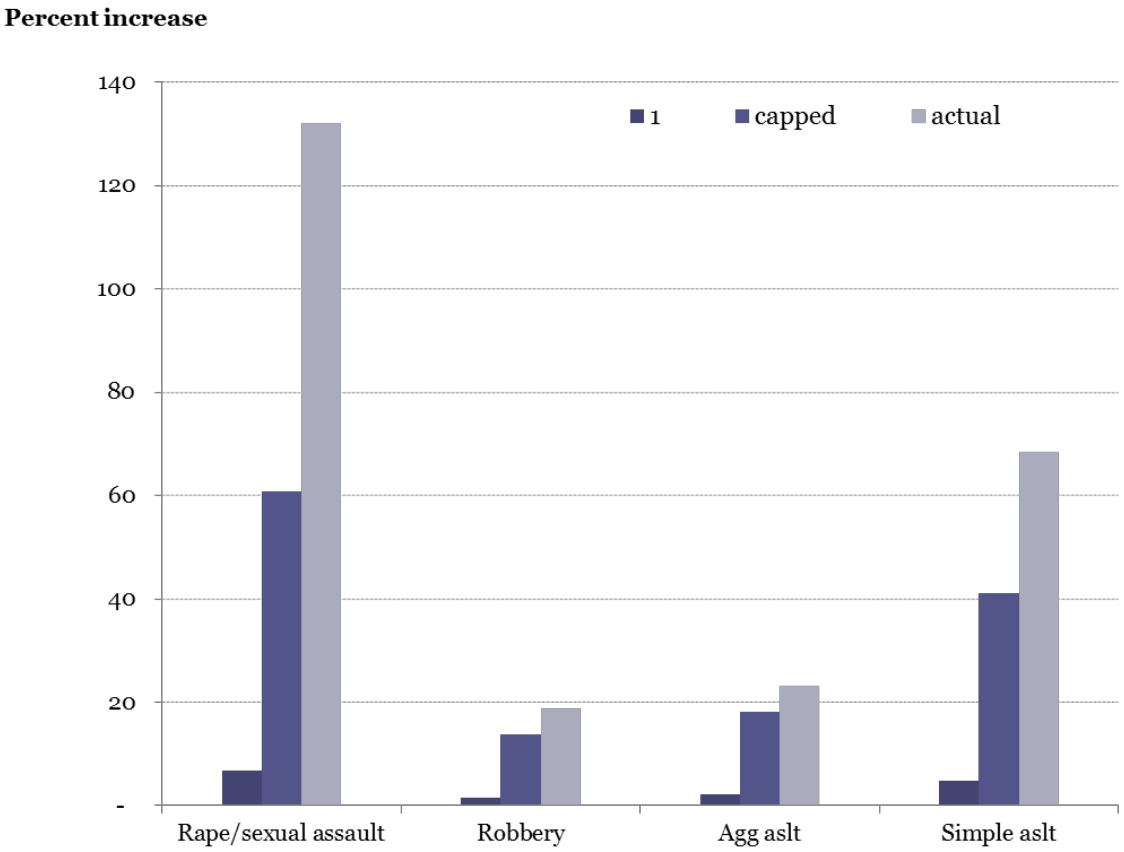


Figure 10. Crime type, coefficient of variation, 2012

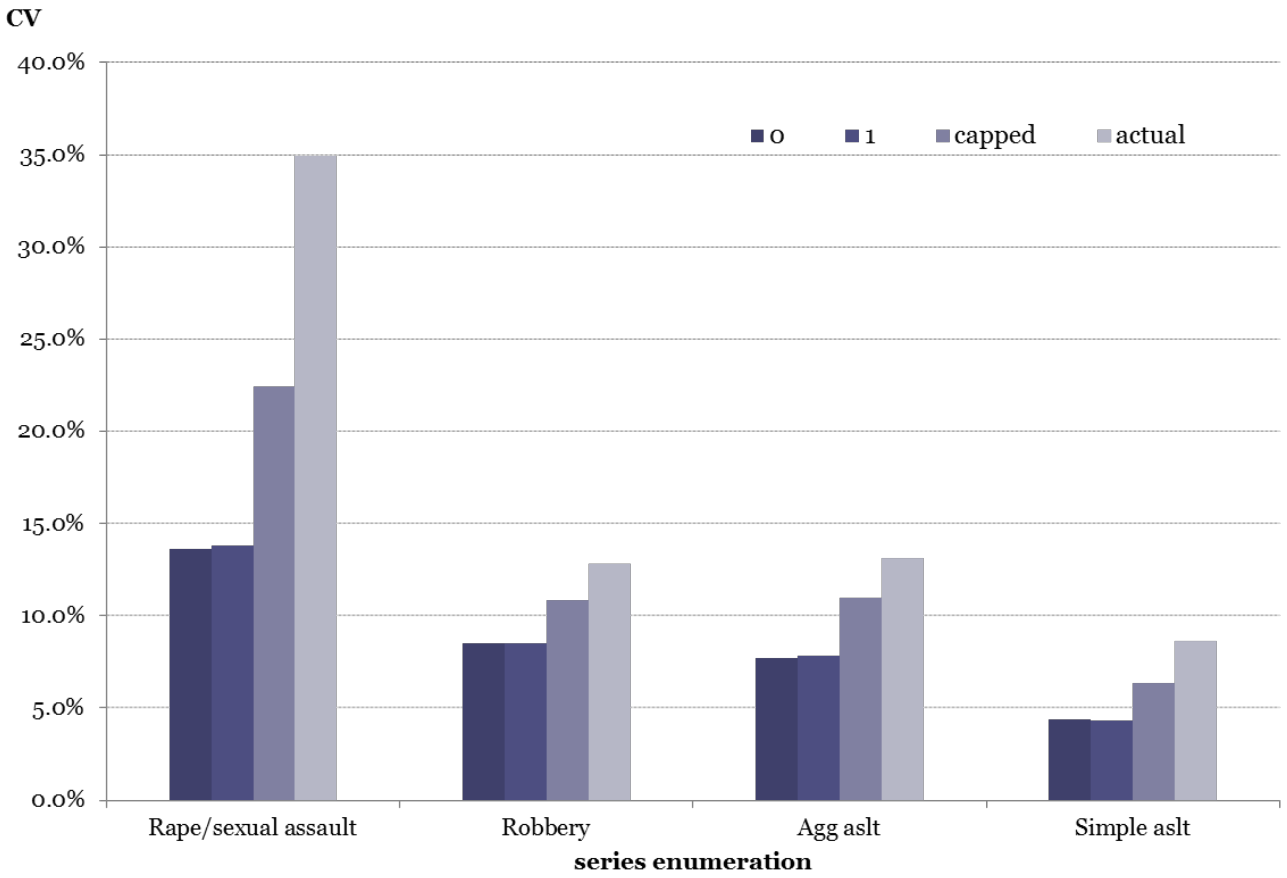


Table 3. Enumeration options: Impact of series counting on point estimate, standard errors, and coefficient of variation (CVs), NCVS 1993, 2012

			series enumeration scheme				Change in estimates to 0			Change in estimates to 1	
Total violent crime			**0**	**1**	**capped**	**actual**	**1**	**capped**	**actual**	**capped**	**actual**
1993	Estimate		10,531,582	11,283,987	16,822,618	28,940,198	7.1%	59.7%	174.8%	49.1%	156.5%
	Standard Error		198,282	206,109	403,246	1,417,798	3.9	103.4	615.0	95.6	587.9
	95% Confidence Interval	Lower	10,140,067	10,877,017	16,026,395	26,140,707	7.3	58.1	157.8	47.3	140.3
		Upper	10,923,097	11,690,957	17,618,841	31,739,688	7.0	61.3	190.6	50.7	171.5
	Coefficient of Variation		.019	.018	.024	.049	-3.0	27.3	160.2	31.2	168.2
	Unweighted Count		3342	3593	3593	3593	7.5	7.5	7.5	0.0	0.0
2012	Estimate		5,081,382	5,283,329	6,842,593	7,988,383	4.0	34.7	57.2	29.5	51.2
	Standard Error		199,234	208,971	376,429	590,922	4.9	88.9	196.6	80.1	182.8
	95% Confidence Interval	Lower	4,687,988	4,870,709	6,099,321	6,821,587	3.9	30.1	45.5	25.2	40.1
		Upper	5,474,776	5,695,949	7,585,866	9,155,178	4.0	38.6	67.2	33.2	60.7
	Coefficient of Variation		.039	.040	.055	.074	0.9	40.3	88.7	39.1	87.0
	Unweighted Count		1596	1661	1661	1661	4.1	4.1	4.1	0.0	0.0

Table 4a. Enumeration options: Impact of series counting on point estimate, standard errors, and coefficient of variation (CVs), NCVS 1993

			series enumeration scheme				Change in estimates to 0			Change in estimates to 1	
Type of Crime			**0**	**1**	**capped**	**actual**	**1**	**capped**	**actual**	**capped**	**actual**
Rape & Sexual Assault	Estimate		521,223	565,153	898,239	1,700,490	8.4	72.3	226.2	58.9	200.9
	Standard Error		27,327	29,729	85,337	534,980	8.8	212.3	1857.7	187.0	1699.5
	95% Confidence Interval	Lower	467,264	506,452	729,738	644,152	8.4	56.2	37.9	44.1	27.2
		Upper	575,182	623,854	1,066,740	2,756,827	8.5	85.5	379.3	71.0	341.9
	Coefficient of Variation		.052	.053	.095	.315	0.3	81.2	500.1	80.6	498.1
	Unweighted Count		160	177	177	177	10.6	10.6	10.6	0.0	0.0
Robbery	Estimate		1,268,704	1,325,480	1,752,667	2,406,190	4.5	38.1	89.7	32.2	81.5
	Standard Error		51,364	52,152	104,109	175,942	1.5	102.7	242.5	99.6	237.4
	95% Confidence Interval	Lower	1,167,283	1,222,504	1,547,099	2,058,785	4.7	32.5	76.4	26.6	68.4
		Upper	1,370,125	1,428,455	1,958,235	2,753,594	4.3	42.9	101.0	37.1	92.8
	Coefficient of Variation		.040	.039	.059	.073	-2.8	46.7	80.6	51.0	85.8
	Unweighted Count		390	410	410	410	5.1	5.1	5.1	0.0	0.0
Aggravated Assault	Estimate		2,532,303	2,650,041	3,481,055	4,907,670	4.6	37.5	93.8	31.4	85.2
	Standard Error		65,479	68,647	165,069	970,942	4.8	152.1	1382.8	140.5	1314.4
	95% Confidence Interval	Lower	2,403,012	2,514,494	3,155,121	2,990,512	4.6	31.3	24.4	25.5	18.9
		Upper	2,661,594	2,785,587	3,806,989	6,824,829	4.7	43.0	156.4	36.7	145.0
	Coefficient of Variation		.026	.026	.047	.198	0.2	83.4	665.1	83.1	663.7
	Unweighted Count		776	815	815	815	5.0	5.0	5.0	0.0	0.0
Simple Assault	Estimate		6,209,352	6,743,313	10,690,657	19,925,848	8.6	72.2	220.9	58.5	195.5
	Standard Error		139,267	144,611	290,345	770,695	3.8	108.5	453.4	100.8	432.9
	95% Confidence Interval	Lower	5,934,364	6,457,774	10,117,361	18,404,083	8.8	70.5	210.1	56.7	185.0
		Upper	6,484,340	7,028,852	11,263,953	21,447,613	8.4	73.7	230.8	60.3	205.1
	Coefficient of Variation		.022	.021	.027	.039	-4.4	21.1	72.4	26.6	80.4
	Unweighted Count		2016	2191	2191	2191	8.7	8.7	8.7	0.0	0.0

Table 4b. Enumeration options: Impact of series counting on point estimate, standard errors, and coefficient of variation (CVs), NCVS 2012

			series enumeration scheme				Change in estimates to 0			Change in estimates to 1	
Type of Crime			**0**	**1**	**capped**	**actual**	**1**	**capped**	**actual**	**capped**	**actual**
Rape & Sexual Assault	Estimate		215,669	230,240	346,830	500,788	6.8	60.8	132.2	50.6	117.5
	Standard Error		29,416	31,789	77,882	174,881	8.1	164.8	494.5	145.0	450.1
	95% Confidence Interval	Lower	157,586	167,472	193,049	155,478	6.3	22.5	-1.3	15.3	-7.2
		Upper	273,753	293,008	500,610	846,097	7.0	82.9	209.1	70.9	188.8
	Coefficient of Variation		.136	.138	.225	.349	1.2	64.6	156.0	62.6	152.9
	Unweighted Count		70	74	74	74	5.7	5.7	5.7	0.0	0.0
Robbery	Estimate		652,000	662,008	741,756	774,332	1.5	13.8	18.8	12.0	17.0
	Standard Error		55,389	56,296	80,468	99,270	1.6	45.3	79.2	42.9	76.3
	95% Confidence Interval	Lower	542,632	550,849	582,869	578,320	1.5	7.4	6.6	5.8	5.0
		Upper	761,368	773,167	900,642	970,344	1.5	18.3	27.4	16.5	25.5
	Coefficient of Variation		.085	.085	.108	.128	0.1	27.7	50.9	27.6	50.8
	Unweighted Count		197	201	201	201	2.0	2.0	2.0	0.0	0.0
Aggravated Assault	Estimate		842,728	860,240	996,106	1,037,449	2.1	18.2	23.1	15.8	20.6
	Standard Error		64,752	67,510	109,362	136,072	4.3	68.9	110.1	62.0	101.6
	95% Confidence Interval	Lower	714,873	726,938	780,168	768,771	1.7	9.1	7.5	7.3	5.8
		Upper	970,584	993,542	1,212,045	1,306,128	2.4	24.9	34.6	22.0	31.5
	Coefficient of Variation		.077	.078	.110	.131	2.1	42.9	70.7	39.9	67.1
	Unweighted Count		261	266	266	266	1.9	1.9	1.9	0.0	0.0
Simple Assault	Estimate		3,370,984	3,530,841	4,757,902	5,675,813	4.7	41.1	68.4	34.8	60.7
	Standard Error		147,281	153,589	302,323	490,379	4.3	105.3	233.0	96.8	219.3
	95% Confidence Interval	Lower	3,080,173	3,227,574	4,160,954	4,707,543	4.8	35.1	52.8	28.9	45.9
		Upper	3,661,795	3,834,108	5,354,849	6,644,083	4.7	46.2	81.4	39.7	73.3
	Coefficient of Variation		.044	.043	.064	.086	-0.4	45.4	97.7	46.1	98.6
	Unweighted Count		1068	1120	1120	1120	4.9	4.9	4.9	0.0	0.0