Comparing the Selection of All Household Members to a Subsample of Household Members: Can Less be More?

Vincent G. Iannacchione[1] and **Bonnie E. Shook-Sa**[2]
[1]RTI International, 701 13th ST NW, Suite 750, Washington DC 20005
[2]RTI International, PO Box 12194, Research Triangle Park, NC 27709

Abstract

The decision to select a subsample of eligible members of a sampled household (HH) is influenced by a number of factors including burden on the HH, data quality, cost, and the sampling variance of survey estimates. Design effects quantify the influence of a complex sampling design on the variance of survey estimates. Selecting a subsample of eligible persons within a sampled HH can have counteracting impacts on design effects. On one hand, subsampling in multi-person HHs increases the design effects attributable to unequal weighting. Conversely, subsampling could reduce the design effects attributable to clustering because the potential intra-household correlation among respondents in the same HH is reduced or eliminated. If the reduction in correlation is greater than the increase caused by unequal weighting, subsampling can achieve the same sampling variance as selecting all eligible HH members with less cost and burden. We present the results of a simulation study that evaluates the design effects associated with subsampling HH members on personal victimization rates based on the 2008 National Crime Victimization Survey which selected all persons 12 and older in a sampled HH.

Key Words: design effects, intra-household correlation, unequal weighting

Introduction

The National Crime Victimization Survey (NCVS) is a survey of the U.S. civilian non-institutionalized population that focuses on personal and property crimes. Within each selected household, a screening questionnaire is administered to all persons aged 12 and older to determine whether they were victims of personal crimes during the previous six-months. Personal crimes are those committed against individuals and include rape/sexual assault, robbery, assault, and personal theft. In addition, a household screener is administered to a single household member (the household respondent) and is used to report property crimes against the household, such as burglary, theft, and motor vehicle theft. The household respondent is the household member chosen by the interviewer who appears to be the most knowledgeable about the household composition and is able to answer the household screener questions dealing with crimes against the household (U. S. Department of Justice, 2008). Additional information is collected about the details of each incident of personal or household victimization reported in the screener. These data are used to estimate yearly victimization rates (VRs) and changes in VRs from year to year. Additional information about calculating VRs is included in Section 3.

The NCVS is a panel survey in which each sampled household (HH)[1] or household equivalent[2] is interviewed once every 6 months over a 3-year period for a total of seven interviews. Currently,

[1] A household is one or more rooms occupied as separate living quarters or intended for occupancy as separate living quarters. The occupants must have direct access from outside the building or through a common hall or entry.

everyone 12 years of age or older in a sampled HH is asked if they were victims of crimes that occurred during the previous 6 months. In response to a recommendation from the National Research Council (2008), the Bureau of Justice Statistics, which administers the NCVS, is considering restricting the sample to one eligible person from each sampled HH. This change in the NCVS sample design could have far-reaching and unknown implications for response rates, survey costs, and survey estimates and their associated sampling variances.

From a data quality standpoint, a one-person per HH design is preferable to either the current 'all person' design or a two-person per HH design because they may be subject to certain biases for intra-familial crimes such as domestic violence. For example, under the current design a husband who has been interviewed may tell his wife not to report domestic violence. The resulting underreporting could be reduced with a single respondent per HH design (assuming privacy is maintained in the interview setting) if the wife is selected because she would be the only person interviewed from the HH. These potential biases cannot be measured in a simulation study. Therefore, this analysis is restricted to the effects of subsampling eligible persons within HHs on the precision of resulting survey estimates.

In a complex sample design like the one used for the NCVS, a number of factors affect the sampling variance of a survey estimate. These include the effects of stratification, clustering, and unequal weighting caused by differential selection and response rates. The net effect these factors have on sampling variances can be measured with a *design effect* or *deff,* which is the ratio of the actual (i.e., design-consistent) variance of an estimate to the variance of a simple random sample of the same size (Kish, 1965). Design effects quantify how the complex design of the sample affects the precision of survey estimates. The larger the design effect, the bigger the sample size required to maintain the desired level of precision.

Restricting the sample to a subsample of eligible persons per HH could have counteracting impacts on the design effects of estimates related to personal VRs. On one hand, the design effects would increase as a result of unequal selection probabilities associated with the sub-selection of persons from multi-person HHs. Conversely, sub-selecting persons from each sampled HH might reduce the design effects because the potential correlation among crimes reported by members of the same HH is reduced. Note that property crimes would not be affected by one person per HH sampling because they are reported by a single household respondent within each HH.

We present the results of a simulation study based on the 2008 NCVS public-use database. The specific research question addressed by the simulation is the following:

> *How does subsampling within NCVS households affect the design effects associated with personal victimization rates?*

To answer this question, we focus on the impact of within-household subsampling on the precision of survey estimates. We begin with a brief description of the current NCVS sampling and weighting methodology. We then describe the selection of the simulation samples that we used to estimate the VRs and the design effects that may accompany either a one-person or two-person per HH selection strategy. We estimate the number of HHs that would be needed to equalize the precision of the current 'all person' sample design and that of a one-person or two-person design. We conclude by discussing the advantages and disadvantages of within-household subsampling.

---

[2] Group quarters are living quarters where residents share common facilities or receive formally authorized care. For the NCVS, group quarters are divided into clusters of four expected persons. These clusters are referred to as household equivalents.

Summary of the NCVS Sampling Design

The NCVS uses a stratified, four-stage sampling design to estimate crime victimization rates for the national civilian non-institutionalized population age 12 and older (U.S. Census Bureau, 2009). At the first stage, Primary Sampling Units (PSUs) are demographic areas consisting of large metropolitan areas, counties, or groups of adjacent counties. Large PSUs are included in the sample automatically and each is assigned its own stratum. These PSUs are considered to be self-representing because all of them are selected. The remaining PSUs, called non-self-representing, are combined into strata by grouping PSUs with similar geographic and demographic characteristics, as determined by the decennial census used to design the sample.

At the second stage, each selected PSU is divided into segments (clusters of about four HHs each), and a systematic sample of segments is selected. At the third stage, all HHs in a sampled segment are selected, and at the fourth stage, all persons age 12 and older are selected from each sampled HH. This type of sampling design enables the selection of a self-weighting probability sample of eligible persons. That is, prior to any weighting adjustments for nonresponse or noncoverage, each eligible person has the same design weight which is the inverse of the overall probability of being selected.

Because of the complex sampling design used for the NCVS*,* the usual sample variance that assumes simple random sampling needs to be multiplied by the design effect to approximate the sample variance associated with the complex design. The design effect is a useful metric to examine when comparing alternative sampling designs. In general, the overall design effect ($DEFF_T$) may be approximated as the product of two components: $DEFF_C$, which is attributable to clustering, and $DEFF_W$, which is attributable to differential sampling rates (or unequal weighting). That is,

$$DEFF_T = DEFF_C * DEFF_W .$$

In a four-stage design used for the NCVS, the design effect attributable to clustering can be approximated as:

where is the average number of sampled persons per PSU, and ρ_1 is the intracluster correlation that measures the homogeneity of the characteristic being measured for persons within the PSUs. Similarly, is the average number of sampled persons per segment, and ρ_2 is the intracluster correlation for persons within segments. Finally, is the average number of sampled persons per household, and ρ_3 is the intracluster correlation for persons within households.[3]

The design effect attributable to differential sampling rates and weighting adjustments for nonresponse or noncoverage (Kish, 1965) can be expressed as:

where W_i is the analysis weight assigned to respondent *i*.

Within-household subsampling would cause $DEFF_C$ to be smaller than that associated with the current NCVS design because the third component ( ) would either be reduced with a two-person per HH design or eliminated with a one-person per HH design. Conversely, $DEFF_W$ would be greater because of the unequal weighting caused by the selection of a subsample from

[3] See Hanson, Hurwitz, and Madow (1953, p. 401) for more details.

each multi-person HH. The combined effect on $DEFF_T$ would depend on the relative decrease of $DEFF_C$ to the increase in $DEFF_W$. By simulating the selection of a subsample of respondents from each multi-respondent HH, the combined effects of $DEFF_C$ and $DEFF_W$ on $DEFF_T$ can be estimated for personal victimization rates.

Selection of the Simulation Samples

The 2008 NCVS public-use database contains survey data for 88,700 respondents in 48,111 unique HHs for the 2008 calendar year (U. S. Department of Justice, 2008). A respondent is a person who completed the screening interview during one or two quarters in 2008. Among these, 69,007 respondents (77.8%) are in 28,418 HHs with 2 or more respondents. The remaining 19,693 respondents (22.2%) are in single-respondent HHs. HHs with multiple respondents had between 2 and 12 respondents per HH. Selecting one respondent from each of the 28,418 multi-person HHs would reduce the total number of respondents to 48,111 (19,693+28,418), which is 54.2 percent of the 88,700 NCVS respondents in 2008. Selecting two respondents from each of the 8,211 HHs with three or more respondents would reduce the total number of respondents to 76,529 (86 percent of the 88,700 NCVS respondents in 2008).

The distribution of respondents per HH is shown in Table 1. Compared to single-respondent HHs, multi-respondent HHs tended to have higher concentrations of male respondents (49% to 42%), Hispanic respondents (14% to 9%), and respondents aged 12 to 24 (22% to 8%). To account for the demographic fluctuations associated with subsampling within a HH, we selected two sets of 1,000 replicated samples from the NCVS public-use database for the 2008 collection year. For the one-person simulation, we selected one respondent at random from each of the 28,418 HHs with 2 or more respondents. The one-person per HH replicate samples consisted of all respondents from single-respondent households and one randomly selected respondent from multi-respondent HHs. For the two-person simulation, we selected two respondents from each of 8,211 HHs with 3 or more respondents. The two-person per HH replicate samples consisted of all respondents from single- and two-respondent households and two randomly selected respondents from HHs with three or more respondents. Although some respondents completed the screening interview twice during 2008, each respondent's probability of selection was independent of the number of times he/she responded.

Table 1. Distribution of Responding Households and Interview Respondents in the 2008 NCVS by Number of Respondents per Household

Number of Interview Respondents per Household	Responding Households[1]		Interview Respondents[2]	
	#	%	#	%
One	19,693	41%	19,693	22%
Two	20,207	42%	40,414	46%
Three or More	8,211	17%	28,593	32%
Overall	**48,111**	**100%**	**88,700**	**100%**

[1] NCVS households with one or more interview respondents during 2008.

[2] Persons who completed a NCVS screening interviews during one or more quarters in 2008.

The NCVS public use files include both person- and incident- level data. Both files are needed to calculate victimization rates. The person-level dataset contains one record for each screening interview while the incident-level dataset contains one record for each criminal incident reported. Personal victimization rates are the estimated number of criminal victimizations per 1,000 persons. The VR for crime *C* in domain *D* is calculated as

$$VR_{C,D} = \frac{\sum_{k \varepsilon C,D} WGTVICCY_k}{\sum_{i \varepsilon D} WGTPERCY_i} * 1000$$

where $\sum_{k \varepsilon C,D} WGTVICCY_k$ is the estimated number of criminal victimizations for crime *C* in domain *D*, calculated by summing the victimization weights (*WGTVICCY*) from the incident-level file for all reported incidents of crime *C* by persons in domain *D*.

$\sum_{i \varepsilon D} WGTPERCY_i$ is the estimated number of persons in domain *D*, which is calculated by summing the person-level weights (*WGTPERCY*) from the person-level file for all persons in domain *D*.

We extracted all of the 2008 collection year interviews and crime incidents for all respondents selected during subsampling. We then applied two adjustment factors to the person weights of respondents in multi-respondent HHs to account for within-HH subsampling. First, we adjusted the person-level weights to reflect the respondents' within-HH probabilities of selection. We then post-stratified the weights of respondents in HHs where subsampling occurred to full-sample totals for respondents in multi-respondent HHs for the 42 domains defined by crossing gender (Male, Female), race (White, Black, Other), and age category (12–15, 16–19, 20–24, 25–34, 35–49, 50–64, 65+). The weights for persons residing in HHs where subsampling was not implemented were not adjusted.

After adjusting the person-level weights for subsampling, we adjusted the weights of victimizations on the incident file by multiplying the victimization weight by the ratio of the new person weight to the original person weight. The subsampling and weighting process is outlined for the one-person subsample in Figure 1. An analogous process was used for the two-person subsample.

Results

We calculated VRs for the one- and two-person per HH replicate samples using the adjusted victimization weights to estimate the total number of victimizations in the numerator and the estimated number of persons in the denominator. We calculated VRs and their associated sample variances for personal crimes (rape/sexual assault, robbery, assault, and personal theft). VRs for the full sample were calculated using the original victimization weights and person weights for respondents from single and multi-respondent HHs. Overall VRs for the one and two person per HH samples were obtained by averaging across the VRs obtained in the 1,000 replicate samples. The resulting design effects of VRs were used to compare the precision of estimates obtained from the one- and two- person designs to estimates calculated under the existing design.

We generated design effects for VRs for the full sample and each of the replicate one- and two-person per HH samples for rape/sexual assault, robbery, assault, and personal theft for domains of interest. Variance estimates were computed using the SUDAAN software (RTI, 2008) with the pseudo-stratum code and the half-sample code as described in the 2008 NCVS codebook. We obtained overall design effects for the one- and two-person samples by calculating the mean design effects across the 1,000 replicate samples. The median design effects across the four crimes are shown in Table 2.

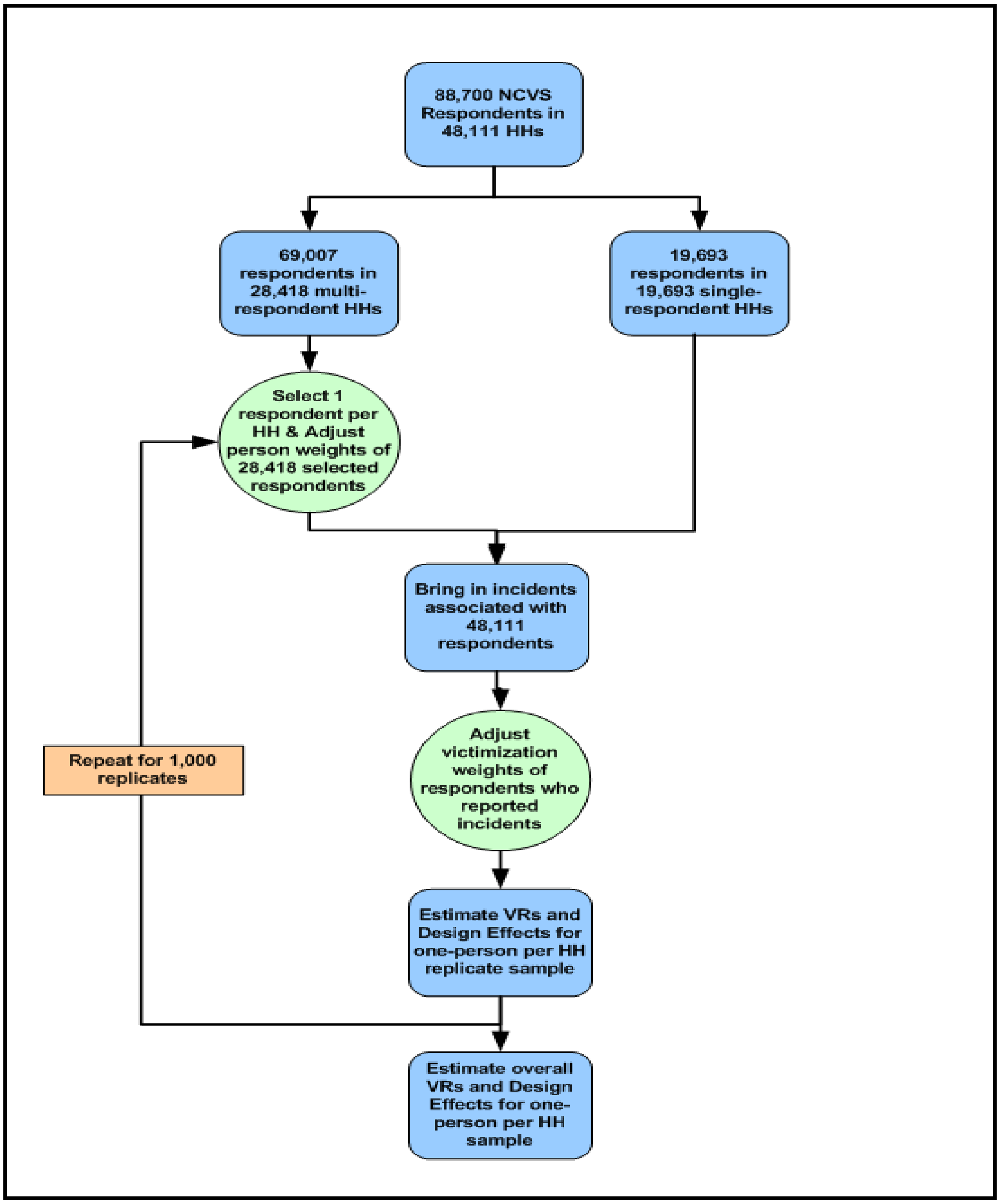

Figure 1. Overview of the sampling, weighting, and estimation process for the one-person per HH replicate samples

Table 2. Median Design Effects for Victimization Rates[1]:
Full Sample vs One-Person vs Two-Persons per Household Samples

	$DEFF_C$[2]			$DEFF_W$[3]			$DEFF_T$[4]		
Domain	**Full Sample**	**Two Persons[5]**	**One Person[5]**	**Full Sample**	**Two Persons[5]**	**One Person[5]**	**Full Sample**	**Two Persons[5]**	**One Person[5]**
Overall	1.26	1.23	1.16	1.09	1.24	1.53	1.38	1.52	1.77
Gender									
Male	1.33	1.30	1.21	1.09	1.23	1.50	1.45	1.61	1.82
Female	1.19	1.13	1.07	1.10	1.24	1.55	1.31	1.39	1.66
Race									
White Only	1.19	1.12	1.05	1.08	1.23	1.52	1.28	1.38	1.59
Black Only	1.39	1.27	1.10	1.14	1.27	1.61	1.58	1.61	1.78
Other	1.20	1.09	1.06	1.10	1.20	1.43	1.31	1.31	1.51
Hispanic Origin									
Hispanic	1.02	0.95	0.94	1.08	1.27	1.56	1.11	1.21	1.46
Non-Hispanic[6]	1.22	1.23	1.17	1.09	1.23	1.52	1.34	1.51	1.79
Age									
12-15	1.22	1.21	1.17	1.07	1.14	1.14	1.30	1.38	1.34
16-19	1.17	1.11	1.06	1.05	1.20	1.34	1.24	1.33	1.43
20-24	1.16	1.04	0.94	1.10	1.23	1.48	1.27	1.28	1.39
25-34	1.30	1.21	1.17	1.10	1.18	1.43	1.43	1.43	1.67
35-49	0.99	0.93	0.88	1.05	1.15	1.40	1.04	1.08	1.23
50-64	1.09	1.07	1.00	1.06	1.11	1.33	1.16	1.19	1.32
65+	1.16	1.13	1.15	1.05	1.08	1.21	1.22	1.23	1.39

[1] Median design effects for victimization rates associated with rape/sexual assault, robbery, assault, and personal theft.
[2] Design effect attrbutable to clustering.
[3] Design effect attributable to unequal weighting.
[4] Overall design effect is the product of *DEFFc* and *DEFFw* .
[5] Median design effects for the one and two-person per household samples are averaged across the 1,000 replicate samples.
[6] Includes 184 persons with unknown Hispanicity.

As expected, the median $DEFF_C$ is lower for the one- and two-person samples than for the full sample because the one-person per HH sample eliminates the intra-household correlation and the two-person per HH sample reduces it. However, the opposite is true for $DEFF_W$. As Figure 2 shows, the unequal probabilities of selection for persons in multi-respondent HHs causes $DEFF_W$ to be highest for the one-person sample, the next highest for the two-person sample, and the lowest for the full sample. When combined, the loss in precision attributable to unequal weighting outweighs the gains in precision from eliminating or reducing within-household clustering. The increased total design effects ($DEFF_T$) causes a loss in precision when estimating VRs with either the one- or two-person per HH subsampling.

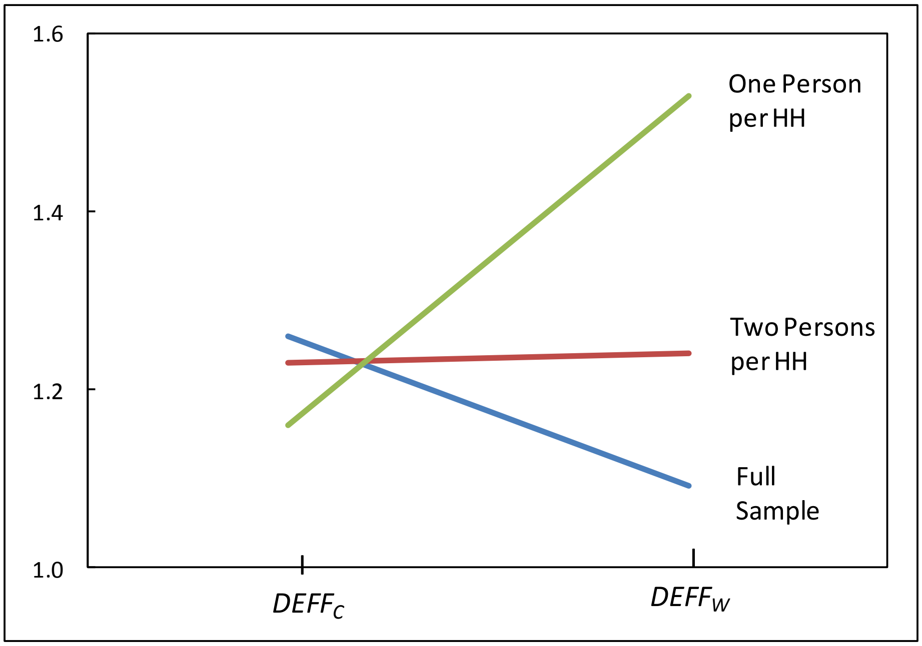

Figure 2. Comparison of median design effects attributable to clustering (*$DEFF_C$*) and unequal weighting (*$DEFF_W$*) for the full sample and the within-household subsamples

To determine the stability of the one- and two-person per HH design effects, we calculated the simulation variance and coefficient of variation (CV) for each estimate. The simulation variance and CV of a design effect are defined as:

$$Var(deff) = \frac{1}{R-1}\sum_{r=1}^{R}(deff_r - \overline{deff})^2,$$

where $\overline{deff} = \frac{1}{R}\sum_{r=1}^{R} deff_r$ and R is the number of simulation samples (R = 1,000), and

$$CV(deff) = \frac{\sqrt{Var(deff)}}{\overline{deff}}.$$

The CVs associated with the three types of design effects are presented in Table 3 for both the one-and two-person per HH designs. The table shows that $DEFF_C$ is more variable than $DEFF_W$ and accounts for most of the variability in the overall design effect. The stability of the CVs indicates that our conclusions about increased design effects are not subject to excessive random variation.

Table 3. Coefficients of Variation for Median Design Effects[1] Associated with Victimization Rates

	One Person per HH	Two Persons per HH
$DEFF_C$ [2]	0.20	0.11
$DEFF_W$ [3]	0.02	0.01
$DEFF_T$ [4]	0.20	0.11

[1] Median design effects for victimization rates associated with rape/sexual assault, robbery, assault, and personal theft.
[2] Design effect attrbutable to clustering.
[3] Design effect attributable to unequal weighting.
[4] Product of $DEFFc$ and $DEFFw$.

Discussion

The results of the simulations indicate that subsampling either one or two eligible persons from each multi-person HH selected for the 2008 NCVS is likely to significantly increase the design effects of the victimization rates (VRs). Increased design effects could lead either to increased costs associated with sampling more HHs to maintain the current precision of VR estimates or to a loss in precision of estimates within domains of interest.

The simulated selection of one respondent from each of the 28,418 multi-respondent HHs that participated in the 2008 NCVS reduced the total number of respondents to 48,111 which is 54.2 percent of the 88,700 NCVS respondents in 2008. The full sample and one-person per HH nominal sample sizes could be equalized by enrolling an additional sample of 88,700-48,111 = 40,589 HHs. Simply equalizing the nominal sample sizes, however, does not consider the increased design effects that are associated with a one- or two- person per HH sample.

Using the one-person per HH approach to achieve the same precision as the 2008 full-sample VR estimates for each of the four types of personal crimes would require one respondent to be enrolled from each of 123,898 HHs. This represents an increase of nearly 75,800 participating HHs over the 48,111 HHs achieved in the 2008 NCVS. A two-person per HH design would require at least 18,796 additional participating HHs to equalize precision.

Our results indicating increased design effects with a one-person design are at odds with those reported by Groves and Heeringa (2006). Their empirical study compared the relative sampling variance associated with selecting one adult in a two-person household to selecting both adults and was conducted as part of the National Comorbidity Survey–Replication (NCS-R), which was a national area probability survey designed to measure the prevalence and severity of mental health disorders in the U.S. household population. The study found that the addition of a second adult respondent in eligible two-person households increased the average sampling variance associated with prevalence rates for mental health diagnoses by 10 to 15 percent. In other words, selecting both persons from a two-person household caused the design effects attributable to intra-household correlation to exceed the reduction in design effects attributable to unequal weighting.

The conclusions of the NCS-R empirical study only apply to a population with two adults per household. As a result, single-adult households, which account for approximately 22 percent of both the NCVS and the NCS-R target populations, are excluded. In general, persons living in single-person households will have a much higher selection probability than those living in multi-person households. For example, a person living in a 2-person household will have a within-household selection probability of 0.5, while a person living in a single-person household will be selected with certainty. Combining data from single-person households with data obtained from multi-person households where subselection was implemented increases the design effects attributable to unequal weighting. This increase in unequal weighting can exceed the reduction in design effects caused by the removal of intra-household correlation, resulting in an increase in the total design effects.

Our simulation study indicates that the unequal weighting that results when multi-person HHs are combined with one-person HHs more than offsets any reduction in design effects caused by the lack of intra-household correlation in a one-person per HH selection. However, two important caveats are associated with this analysis.

1. The simulation assumes that the response propensities of NCVS sample members are not significantly affected by within-household subsampling. However, the survey literature suggests that the size of the survey request (intention to interview everyone 12 or older in a HH vs. a subsample) may affect response rates (i.e., the greater the burden, the lower the participation rate). In addition, attempting to interview everyone in a HH may result in privacy concerns that cause deliberate concealment of one or more HH members (Valentine & Valentine, 1971). In addition, a positive (or negative) interview experience for one HH member may help to gain (or discourage) the cooperation of the other HH members. This group dynamic would not apply to a single-respondent design.

2. The cost savings associated with interviewing a subsample of persons in a multi-person HH are not considered. Presumably, these cost savings would offset at least part of the increased cost needed to enroll enough additional HHs to equalize the precision of a within-HH subsample with that of the full sample.

Despite these limitations, this research provides an estimate of the loss in statistical precision that would result if the NCVS were to transition to selecting one or two persons per HH. Although within-household subsampling would reduce the burden on individual HHs, the resulting increase in design effects would lead either to higher costs associated with selecting significantly more HHs or to a loss in statistical precision of NCVS survey estimates.

References

Groves, R. M., & Heeringa, S. G. (2006). Responsive design for household surveys: Tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society A*, 169, part3.

Hanson, M., Hurwitz, W., & Madow, W. (1953). *Sample survey methods and theory, Volume I.* New York: John Wiley & Sons.

Kish, L. (1965). *Survey sampling.* New York: John Wiley & Sons.

National Research Council. (2008). *Surveying victims: Options for conducting the National Crime Victimization Survey.* Panel to review the programs of the Bureau of Justice Statistics. Robert M. Groves and Daniel L. Cork, eds. Committee on National Statistics and Committee on Law and Justice, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Research Triangle Institute (2008). *SUDAAN language manual, release 10.0*. Research Triangle Park, NC: Research Triangle Institute.

United States Census Bureau. (2009). *General description of the 2000 sample design of the National Crime Victimization Survey* (NCVS00SDE01). Internal Census Bureau memorandum from Kathlene Garland, October 22, 2009.

United States Department of Justice, Bureau of Justice Statistics. (2008). *National Crime Victimization Survey, 2008* [Collection Year Record-Type Files] [Computer file]. ICPSR25461-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2009-09-11. doi:10.3886/ICPSR25461

Valentine, C. A., & Valentine, B. L. (1971). *Missing men: A comparative methodological study of underenumeration and related problems*. Washington, DC: U.S. Bureau of the Census.