Journal of Quantitative Criminology, Vol. 6, No. 2, 1990

Using Network Sampling in Crime Victimization Surveys

Ronald Czaja[1,3] and Johnny Blair[2]

Since crime victimizations are statistically rare events, surveys to estimate rates of victimization are difficult and expensive. In this paper, we examine the advantages of network sampling over traditional methods for conducting crime victimization surveys. Network sampling links population households in specified ways, for reporting purposes, in order to increase the probabilities of locating households with particular characteristics. We conducted a reverse record check field experiment to test whether a telephone survey using network sampling is feasible to collect crime victimization data. Three types of crimes—burglary, robbery, and assault—were tested along with two types of networks—relatives and co-workers/close friends. This paper examines the extent to which victims report their victimization experiences in a general crime and victimization interview and the extent to which a randomly selected relative or close friend will report the same victimization incident in an identical interview. A number of multiplicity counting rules are compared in terms of reporting errors and a mean square error analysis.

KEY WORDS: telephone victimization survey; network sampling; reverse record check; mean square error analysis.

1. INTRODUCTION

1.1. Victimization Surveys

Conducting a crime victimization survey is quite difficult and costly because victimizations in any given year are rare. For example, U.S. crime data for 1986 (U.S. Federal Bureau of Investigation, 1987) indicated that there were 225 robberies per 100,000 persons 18 years of age and older, 346 aggravated assaults per 100,000 persons in the same age range, and

[1]Department of Sociology, North Carolina State University, Raleigh, North Carolina.
[2]Survey Research Center, University of Maryland, College Park, Maryland.
[3]To whom correspondence should be addressed at Department of Sociology, North Carolina State University, Box 8107, Raleigh, North Carolina, 27695-8107.

0748-4518/90/0600-0185$06.00/0

1345 burglaries per 100,000 households. In terms of percentages, robberies were 0.2%, assaults were 0.35%, and burglaries were 1.35%. These figures represent events reported to law enforcement agencies. Even if reported crimes are one-sixth to one-half of all crimes committed, specific types of crimes are still rare events.

Garofalo (1977) presents the results of a hypothetical victimization survey using the rates from the eight cities that were included in the High Impact Crime Reduction Program. A sample of 2400 households and 5120 individuals would be expected to yield 11 rapes, 107 robberies, 172 simple and aggravated assaults, 495 personal larcenies, 358 burglaries, and 289 household larcenies. As expected, these data show very low yields for violent crimes (rape, robbery, assault) and would be inadequate for many analyses. Even the nonviolent and household crime data would be inadequate if analysis by demographic groups or by smaller geographical areas was necessary. Consequently, large samples are necessary for good estimates, and such surveys using careful traditional methods are very expensive (Fienberg, 1978).

The problem is exacerbated for local communities that desire to obtain estimates of particular types of crimes. In many instances, local planners need victimization data in considerable geographic detail, such as at the neighborhood level. However, conducting a victimization survey with conventional methods requires resources beyond the capabilities of the local agency or government.

1.2. Network Sampling

One method that has shown encouraging results in the sampling of rare populations is network sampling (Sudman *et al.*, 1988). The primary difference between network and conventional surveys is the number of reporting units to which target respondents (victims) are linked. In conventional surveys, each target respondent is linked to one enumeration unit: his place of residence. In network sampling, respondents are linked to, that is, can be reported by, additional enumeration units (Sirken, 1970). This linkage is accomplished by using a multiplicity counting rule.

An individual's network is defined as the set of enumeration units linked to and eligible to report that individual. The most common counting rules link individuals to the households of prespecified relatives. The resulting sample is a probability sample, although all elements do not have the same probability of selection. For example, assume that a robbery victim has within the study community the following eligible relatives who resided in five separate households: parents (living in one household), two brothers (living in two households), one sister, and one daughter. The victim, there-

fore, could be reported in six different households (five relative households plus his own). If another robbery victim has none or one eligible relative residing in a separate household in the community, he could be reported by a total of one or two households, respectively. In data analysis, cases are appropriately weighted to adjust for the different probabilities of selection. The number of households eligible to report the target respondent is elicited in the interview; this number is used to construct the case weight.

A further possible advantage of network sampling concerns the traditional problems of nonreports of victimizations. Since the key characteristic of the method allows reporting of a crime by someone other than the victim, nonreporting by victims is, to some extent, offset. Also, by allowing respondents to report about others as well as about themselves, comparison of incidence by self-report estimates with incidence by multiplicity estimates permits measures of respondent underreporting (Czaja *et al.*, 1986).

A number of network surveys have been conducted over the past decade, primarily in the areas of health and vital statistics. Network surveys have been used to estimate the incidences of neurological conditions (Sirken and Royston, 1977), diabetes (Sirken *et al.*, 1975), cancer patients (Czaja *et al.*, 1986), heroin users (Rittenhouse and Sirken, 1980), births and marriages (Nathan, 1976), and the seriously ill (Sudman and Freeman, 1988) and to locate racial minority members who served in Vietnam (Rothbart *et al.*, 1982). In general, these studies conclude that for rare populations network samples provide estimates with smaller sampling variances when they increase the yield of the target population.

Network sampling does introduce some complications and costs compared with traditional surveys. A key component of surveys with multiplicity is the effect of counting rules on reporting errors: overreporting, underreporting, and incomplete or inaccurate reports. Which counting rules to use in a network survey is a critical decision. On the one hand, the use of a broad counting rule allows the target person or household to be reported by a larger number of respondents, which in theory minimizes the number of households to contact. However, as one expands the number of potential respondents, increases in reporting errors are also likely.

There is very little guidance from the social science literature on how broad or narrow to make the reporting rules for specific research topics. The reporters must know about the event, be willing to report it, and in most cases, have some knowledge of the details of the event. A large part of determining the feasibility of using network sampling is measuring the tradeoff on total survey error between higher response bias and reduced sampling variances. In general, expanding the definition of the eligible network increases response bias but, at the same time, contributes to reduced sampling variances.

Further, network surveys cost slightly more than a conventional survey of the same sample size owing to the additional questions that must be asked to elicit the network reports and network sizes. Although, to date, no research has tested a model that jointly considers cost, bias, and variances, all of these need to be considered, even if informally, in determining the efficacy of network sampling for a particular application.

In summary, for network sampling to be successful, a minimum of two broad conditions must be satisfied.

(1) The network respondents must be able and willing to report about the event or characteristic of interest; and
(2) they must know the size of the eligible reporting network.

For our research, four conditions were relevant:

(1) Network respondents must know about the victimization;
(2) They must be willing to report about it;
(3) They must have a reasonable knowledge of the time period in which it occurred; and
(4) They must know the size of the target person's (victim's) reporting network.

One approach that has been successfully used in the past to identify methodological weaknesses in victimization reports is the record check survey (RCS) (Sparks, 1982), even though reverse record check designs, in themselves, raise a number of methodological issues (Biderman and Lynch, 1981). The purpose of the RCS is to determine whether persons who have reported incidents to the police are willing to report them in a survey interview and, if so, how accurately they report the data and other details of the incident. RCSs are a cost-effective means of validating a survey questionnaire, especially if the survey respondents occur with relative infrequency in the population (Dodge, 1983). In the present study, we conducted a RCS and linked network sampling with telephone interviewing to determine if these two methodologies would be effective for conducting local victimization surveys while also increasing the precision of sample estimates. In experimental comparisons of mode of interview, Groves and Kahn (1979) have shown that the quality of data is not seriously affected in telephone interviewing; these results carry over into the area of victimization surveys (Tuchfarber and Klecka, 1976; Skogan, 1976). In this paper, we compare the results from a conventional survey with the results from network surveys, using a number of different counting rules, in terms of reporting rates, bias, and the mean square error.

2. MULTIPLICITY ESTIMATION

The counting rules that we assess in this paper are the following:

Rule 1: A conventional rule in which the victims are linked only to their usual residence.

Rule 2: A sibling rule in which the victims are linked to their usual residence and to the residences of their siblings.

Rule 3: A parent and children rule in which the victims are linked to their usual residence and to the residences of their parents and children.

Rule 4: A relative rule in which the victims are linked to their usual residence and to the residences of their siblings, parents, and children.

Rule 5: A close friend/co-worker rule in which the victims are linked to their usual residence and to the residences of their close friends.

Rule 6: A combined rule in which the victims are linked to their usual residence and to the residences of their relatives and close friends.

In the following expressions, the multiplicity of an event is referred to as Rule 1, Rule 2, etc. The multiplicity model that we are using was first developed by Sirken (1979) and further elaborated in Casady *et al.* (1985) and Czaja *et al.* (1986). Let N be the size of the population at risk and $\theta = V/N$ be the incidence of victimization for the population at risk. Then, a multiplicity estimator for θ from a simple random sample of m households from a universe of M households for rule r is

$$\hat{\theta} = \frac{(M/m)}{N}\left[\sum_{i=1}^{M} a_i \sum_{j=1}^{V} \beta_{rij}/S_{rj}\right], \qquad r = 1, 2, 3, 4, 5, 6 \tag{1}$$

where

V = the total number of victims in the population at risk

$$a_i = \begin{cases} 1 & \text{if household } i \text{ is selected} \\ 0 & \text{otherwise} \end{cases}$$

$$\beta_{rij} = \begin{cases} 1 & \text{if event } j \text{ is reported at household } i \\ 0 & \text{otherwise} \end{cases}$$

S_{rj} = multiplicity of event j for rule r

For the conventional rule, Rule 1, $S_{1j} = 1$, and for Rules 2–6, $S_{rj} \geq 1$ is the total number of different households in which the victim and the eligible network members reside.

The expectations for the mean, variance, and bias for the estimators are the following:

Conventional counting rule:

$$E(\hat{\theta}_1) = \theta p_1 \tag{2}$$

$$\text{Var}(\hat{\theta}_1) = \frac{\theta p_1}{m}[M/N - (\theta p_1)] \tag{3}$$

$$\text{bias}(\hat{\theta}_1) = -\theta(1 - p_1) \tag{4}$$

Multiplicity counting rule:

$$E(\hat{\theta}_r) = \theta\left[p_r + (p_1 - p_r)\frac{\sum_{j=1}^{V} 1/(S_{rj})}{V}\right], \qquad r = 2, 3, 4, 5, 6 \tag{5}$$

$$\text{var}(\hat{\theta}_r) = \frac{\theta(M/N)}{m}\left[p_r \frac{\sum_{j=1}^{V} 1/(S_{rj})}{V} + (p_1 - p_r)\frac{\sum_{j=1}^{V} 1/(S_{rj})^2}{V}\right] - \frac{[E(\hat{\theta}_r)]^2}{m} \tag{6}$$

$$\text{bias}(\hat{\theta}_r) = -\theta\left[(1 - p_r) - (p_1 - p_r)\frac{\sum_{j=1}^{V} 1/(S_{rj})}{V}\right] \tag{7}$$

where

p_1 = the conditional probability that a victim is reported when his or her residence is selected in the sample

p_r = the conditional probability that the victim is reported at the residence of their relative or close friend as specified in rule r, $r = 2, 3, 4, 5, 6$

$$MSE(\hat{\theta}_r) = \text{var}(\hat{\theta}_r) + \text{bias}^2(\hat{\theta}_r) \tag{8}$$

A number of factors should be noted about the above expressions and the study design. None of the formulas includes finite population correction factors. The variance expression assumes that no household reports more than one crime event. Although in actuality this would not be the case, for the purposes of this feasibility study we were interested only in a report of a single target crime, for which we could validate the details from the police report. Reports of other crimes were irrelevant for our primary purpose.

The expression also assumes that the conditional probability of reporting an event for the combined rules (Rules 4 and 6) are constant across rules. On the latter point, we are assuming this to be true for purposes of simplifying the testing of combined rules even though Table III will show that, as we discuss later, reporting by siblings and children/parents is not constant from household to household. In addition, it should be noted that

these data are not weighted for people who had multiple crime victimizations.

The sample was selected with equal probability from victim reports. People who reported multiple victimizations during the sample period had a slightly higher chance of inclusion in the sample. However, the cost of searching paper lists to determine the total number of reports for each sample member was prohibitive.

3. DESIGN AND IMPLEMENTATION OF THE STUDY

3.1. Design

The logic of the design of this feasibility survey is quite simple. A sample of known crime victims was selected from the city police records of a small Metropolitan Statistical Area (MSA) in Illinois. A few population and housing characteristics of the city are the following. In 1980 the total population exceeded 100,000 persons, of whom 81% were white, 17% were black, and 12% were 65 and over; the median persons/household was 2.46 for whites and 3.18 for blacks; 67% of the occupied housing units were owner occupied; the median income was $23,400 for owner-occupied housing units and $11,200 for renter-occupied units; and approximately 5% of the owner-occupied units and 23% of the renter-occupied units were below poverty level (U.S. Bureau of the Census, 1982a, b, 1983).

A general crime victimization telephone interview was conducted with each victim. In the interview, respondents were asked both about crimes that had happened to them personally and about crimes that had happened to members of their specified network. At the end of the interview, the name and telephone number of a randomly selected member of the victim's defined network were elicited. These network members were called and the same interview was conducted with them. This design permitted simple comparisons of self and network reports. The victim and network member samples were combined with a general population (decoy) sample selected from the same neighborhoods as the victims. The purpose of the decoy sample was to preserve the confidentiality of nonreporting victims and to minimize potential interviewer bias (Turner, 1972; Dodge, 1983).

Several counting rules were examined. One counting rule was a combination of the three closest friends at work or the closest nonwork friends living in the metro area; and the others were rules based on specified relatives of the respondent—parents, siblings, and adult children—living in separate households in the metro area.

A crucial factor in any victimization survey is the ability of respondents to recall victimization events completely and accurately. In a victimization study that employs network sampling, it is important that network members be able and willing to recall the event accurately. Previous research conducted by the Law Enforcement Assistance Administration (LEAA) and the U.S. Bureau of the Census (Kalish, 1974) comparing victimization reports of self-respondents and of proxy respondents within a household has shown that victimization rates for serious crimes were underreported by proxies. Also, and not surprisingly, the details about many of the crimes were unknown or less than complete when reported by proxies (Skogan, 1981).

Our research differs in one important manner. We selected proxy respondents from outside the victims' households. Sirken (1974) and Bradburn and Sudman (1979) have suggested that these types of respondents may be able and willing to provide more accurate information than household members, especially if the event deals with embarrassing or socially undesirable behavior. Other studies (Fischer, 1982; Lopata, 1979) indicate that most persons have social support networks and confidants who are not part of the immediate family. The type and extensiveness of these relationships vary by life cycle, age, sex, and role-relationship variables. In many situations (e.g., teenagers and young adults), friendship networks are more important than family. Thus, the counting rules selected appear to meet the basic criteria necessary for informants reporting in a network sample survey. It is likely that they would know of the event, be willing to report it, and finally, be able to provide the network size of the victims about whom they report (Skogan and Maxfield, 1981).

The selection of types of victimization was guided by three criteria: (i) the victimizations should be of particular interest to criminal justice planners, (ii) they should be ones not easily measured by conventional survey techniques, and (iii) they should be crimes that one would reasonably expect informants to have heard about.

Robbery, burglary, and assault all satisfied these criteria and were feasible in terms of both access to records and the number of events available in a year. Each of these crimes is a rare statistical event. In the counties where we conducted our research, the rate for reported robberies was 122/100,000 persons 18 years of age and older (0.12%), the rate for assaults was 549/100,000 (0.55%), and the rate for burglaries was 2,814/100,000 housing units (2.81%). These rates are low enough to indicate that conducting a victimization survey with conventional methods would require resources beyond the capabilities of the local agency or government.

As criminal offenses, the crimes comprise a high proportion of street crime and are all fear producing, hence making them salient and likely to be known by network members. Robbery as a research subject has the

advantage of being a crime that occurs to a wide range of victims. It is sufficiently underreported in conventional victimization surveys to provide a good comparison with a network survey (Flanagan and McLeod, 1983). Assault disproportionately involves young offenders, as well as offenders known by the victim. Burglary was chosen in order to include a crime against property.

3.2. Procedures

The crime victims were chosen from police department records for the period February through September 1986. The names, addresses, and telephone numbers of network members were, of course, provided by victims with whom interviews were completed. The decoy sample was selected from current telephone directories covering the MSA. Data collection was conducted primarily by telephone, with face-to-face interviewing used for only a few respondents who were not reachable by telephone.

The sample frame for the victimization respondents consisted of two parts. First, we used a tape of 2640 robbery, assault, and burglary cases that occurred in the jurisdiction of the police department from February through September 1986. Second, we used the actual police reports, including the officer's narrative report, for all sample cases. The sample frame provided on tape was sorted by type of victimization, and then systematic random samples were selected. The corresponding police reports were then pulled from the police department files. A decoy sample of 160 telephone numbers was selected. The effective decoy size was larger than this, however, owing both to the nonreports of some of the victim households and to the relatively high percentage of ineligibles.

Once a household was contacted by telephone, respondent selection was conducted in the following manner: After a brief introduction explaining the purpose of the survey as a general community satisfaction study, all household members aged 18 or older were listed, with the sex and age of each person entered on a household chart. A further question was asked about whether anyone else aged 18 or older, such as a friend or another relative, had used the household as a second or temporary residence any time during that year. In this way, we hoped to ensure that all persons aged 18 or older who had resided in the household any time during the year (and hence could have been the crime victim) were listed. For each household contacted, the interviewer was provided with a basic demographic description of the type of respondent (such as female between 20 and 25 years old) that we wanted to interview in that household, if such a person resided there. In this way, we hoped to increase the likelihood of enumerating and interviewing the crime victim without letting the interviewer know

in which households we expected to find victimizations. The same procedure was used for relative, friend, and decoy households.

Once the target respondent was contacted, the interview began with a series of general questions about satisfaction with his/her neighborhood and the area in general. The next section elicited the first names of network members beginning with friends and co-workers and followed by parents, siblings, and adult children living in other households. Only after the names of all network members had been elicited were the questions about victimizations asked. This ensured that respondents would not simply give the names of people who had been victims of crime but would, in fact, give the names of their closest work and nonwork friends. The questions eliciting the various types of victimization were adapted from the National Crime Survey.

The same set of victimization questions was then asked with regard to co-workers, close friends, and relatives. A set of standard demographics about the respondent was then asked. Finally, the interviewer randomly selected one relative and one co-worker or friend. For these two randomly selected individuals, complete name, address, and telephone contact information were asked. When the respondent was unable to give complete information to contact the network member, the name, address, and telephone number of a secondary source who might be able to provide that information were elicited. Interviews were conducted during the period October 1986 through January 1987 by professional interviewers at the Survey Research Laboratory's Telephone Center, University of Illinois. All refusals were reworked.

The time interval between when a victim respondent was interviewed and when the network member was interviewed was approximately 2 weeks. The median was 15 days, and the mean was slightly more than 17 days. We received no indications from interviewers that victim respondents contacted the selected network members and informed them that they might be contacted for an interview.

One area of the data coding was so crucial to the interpretation of the survey results that it is summarized here. Before determining whether the target crime was reported in the interview, three steps had to occur: First, one must have reached the correct household; second, the crime victim or network member must have been enumerated as a household member; and third, that victim or network member must have become the interview respondent, except for the household crime of burglary, where any adult family member who was in residence in the victim household at the time of the crime was eligible.

In order to determine whether the correct victim household was contacted, coders compared the following items from the questionnaire with the

police report: telephone number, age and sex of someone listed in the household chart matching the victim listed in the police report, and number of years of residence at the present address.

Next the coder was to determine whether the victim was correctly enumerated in the household chart. This was done by comparing the following information from the questionnaire with the police report: victim's first name, age of victim within a year, and race of the victim. If the information matched the police report, the coder noted that the victim was enumerated in the household chart.

The final step was to determine whether a crime event was reported and, if so, whether the reported crime event was, in fact, the event of the police report. Because of memory error and other factors, as has been found by others (Miller and Groves, 1985), we did not expect that a target crime reported in the questionnaire would exactly match, in every detail, the same crime in the police report. A series of criteria was developed for each type of crime for purposes of comparison and classification. Whether or not the target crime was reported was classified into four categories: yes, probably yes, probably no, and definitely no. The basis of these classifications was as follows: items such as what was taken, whether the offender was known, whether a weapon was used, and the date of the crime were grouped based on the likelihood that the item would be remembered by the respondent. Matching on all of the items was required for a definite "yes". If there was only a partial match on the key items, the reported date of the incident was used to determine whether there was a "probable," but not certain, match. After coding the data at this level of specificity, it appeared that there was no real difference between the "probably" and the "definitely" categories, so these were combined, producing a simple dichotomy of "yes" or "no."

4. FINDINGS

4.1. Summary of Survey Results

A sample of 688 crime victims was eligible for interviewing (Table I). Of these, 307 were burglary victims, 148 were robbery victims, and 233 were assault victims.

Interviews were obtained with 254 burglary cases (82.7%). Of those interviews, 204 were classified as in-scope and 50 were classified as out-of-scope. A case was dispositioned as in-scope if the correct household was contacted, the respondent was the crime victim according to the police record, and the crime occurred within a designated recall period.

Any of the following reasons resulted in a case being classified as out-of-scope: (i) the respondent was victimized in a commercial location

Table I. Interview Completion Rates for the Victim Sample

Disposition	Burglary	Robbery	Assault	Total	Rate (%)
Eligible sample	307	148	233	688	100.0
Interviews	254	109	196	559	81.3
In-scope	204	69	110	383	—
Out-of-scope	50	40	86	176	—
Refusals	32	14	17	63	9.1
Other (noncontacts or unavailable)	21	25	20	66	9.6

such as a bank or gas station; (ii) the interview was conducted in the wrong household; (iii) the questionnaire items for a few respondents were not appropriate to elicit reporting of the target crime; (iv) the friend/co-worker or relative respondent did not mention the victim as part of their network and; (v) the recall period did not encompass the date of the incident because of a field experiment conducted on telescoping. On this last point, a small portion of the sample was assigned a recall period that did not encompass the date of the target crime. The purpose of this experiment was to investigate forward telescoping by respondents. These respondents are excluded from this analysis.

Table II. Network Completion Rates

Disposition	Burglary	Robbery	Assault	Total	Rate (%)
			Relatives		
Eligible sample	51	33	48	132	100.0
Interviews	41	31	41	113	85.6
In-scope	26	10	16	52	—
Out-of-scope	15	21	25	61	—
Refusals	10	0	3	13	9.8
Other (noncontacts or unavailable)	0	2	4	6	4.6
			Friends/co-workers		
Eligible sample	63	26	38	127	100.0
Interviews	55	23	30	108	85.0
In-scope	21	9	5	35	—
Out-of-scope	34	14	25	73	—
Refusals	7	1	5	13	10.2
Other (noncontacts or unavailable)	1	2	3	6	4.7

Of the 148 robbery cases, interviews were completed with 109 (73.7%)—69 in-scope and 40 out-of-scope. Interviews were completed with 196 assault cases (84.1%)—110 in-scope and 86 out-of-scope.

The network sample included 132 relatives and 127 friends/co-workers (Table II). Of the 132 relative cases, interviews were conducted with 113 (85.6%)—52 in-scope and 61 out-of-scope. Completed interviews were classified as in-scope or out-of-scope using the same criteria that were applied to the victim interviews.

Of the 127 friend/co-worker cases, interviews were conducted with 108 (85.0%)—35 in-scope and 73 out-of-scope.

4.2. Response Model Findings

The reporting rates of victims by their own households and the households of the eligible relative and close friend (work and nonwork) networks are shown in Table III. For the total sample of victim households, 66% of the target crimes were reported, whereas in the network households the rates varied from a low of 26% for siblings to a high of 59% for the child/parent rule. There was significant variability by type of crime, especially for the victim households. Burglary and robbery were reasonably

Table III. Estimated Target Crime Reporting Rates for Victim and Network Households, the Total Sample, and Selected Victim Subdomains

	Reporting households					
Victim subdomain	Victim p_1 (N)	Sibling p_2 (N)	Child, parent p_3 (N)	Sibling, child, parent p_4 (N)	Friend p_5 (N)	Sibling, child, parent, friend p_6 (N)
Total sample	0.66 (383)	0.26 (23)	0.59 (29)	0.44 (52)	0.51 (35)	0.47 (87)
Crime type						
Burglary	0.84 (204)	0.38 (13)	0.62 (13)	0.50 (26)	0.57 (21)	0.53 (47)
Robbery	0.72 (69)	0.20 (5)	0.60 (5)	0.40 (10)	0.67 (9)	0.53 (19)
Assault	0.29 (110)	0.00 (5)	0.55 (11)	0.38 (16)	0.00 (5)	0.29 (21)
Sex						
Male	0.68 (159)	0.20 (10)	0.55 (11)	0.38 (21)	0.57 (14)	0.46 (35)
Female	0.65 (224)	0.31 (13)	0.61 (18)	0.48 (31)	0.48 (21)	0.48 (52)
Race						
White	0.71 (304)	0.25 (20)	0.55 (22)	0.40 (42)	0.50 (32)	0.45 (74)
Nonwhite	0.46 (79)	0.33 (3)	0.71 (7)	0.60 (10)	0.67 (3)	0.62 (13)
Age						
<35	0.60 (167)	0.17 (12)	0.69 (13)	0.44 (25)	0.31 (13)	0.39 (38)
≥35	0.70 (216)	0.36 (11)	0.50 (16)	0.44 (27)	0.64 (22)	0.53 (49)

well reported (84 and 72%, respectively), but assaults were grossly underreported (29%). The results for burglary and robbery are similar to those found in the San Jose reverse record study (Turner, 1972), where 90% of the burglaries and 76% of the robberies were reported. For assault, the results are similar in that it was poorly reported in both studies. Our rates, however, are much lower than the 48% report rate in San Jose.

In general, the victim households were the better reporters, but for a number of subdomains (Table III), one or more of the counting rules had a comparable or higher reporting rate. This occurred for assaults, for victims who were female or nonwhite, and for both age groups. The child/parent rule had the highest reporting rates, with the friend rule being comparable in many instances. Clearly, the sibling rule had the poorest reporting rates.

The expressions for bias are a function of the estimated rate of nonresponse and of multiplicity parameters [see Eq. (7)]. In Table IV we present the ratios of the estimates of the bias for the multiplicity and the conventional counting rules. The pattern of results is essentially the same as for the reporting rates; that is, in most instances the conventional rule has a smaller

Table IV. Ratios of the Bias for Multiplicity Counting Rules to That for the Conventional Counting Rule, by the Total Sample and Selected Victim Subdomains

	Bias ratio = bias $\hat{\theta}_r$/bias $\hat{\theta}_1$				
Victim subdomain	Sibling	Child, parent	Child, sibling, parent	Friend	Sibling, child, parent, friend
Total sample[a]	1.36	1.07	1.29	1.29	1.44
Crime type					
Burglary	1.80	1.41	1.90	2.16	2.52
Robbery	—[b]	—	1.49	—	1.56
Assault	—	0.88	0.94	—	1.01
Sex					
Male	1.44	1.11	1.39	1.23	1.55
Female	1.30	1.03	1.22	1.32	1.37
Race					
White	1.46	1.17	1.47	1.53	1.75
Nonwhite	—	—	0.87	—	0.78
Age					
<35	1.37	0.94	1.18	1.52	1.43
≥35	1.32	1.23	1.40	1.15	1.46

[a] $N = 383$.

[b] Fewer than 10 respondents.

bias than the other rules, and the child/parent rule has the smallest bias ratio among the multiplicity rules. The results are due primarily to the higher reporting rates for these two rules. The pattern is not similar for the sibling rule, where its bias ratio is comparable to the friend rule.

In Table V we present the estimated sample size for which the mean squared error (MSE) of the conventional estimator equals the MSE of the specific multiplicity estimator. The results are shown for an incidence rate of 0.001, for various subdomains, and for a metropolitan area of about 350,000 population. The results are essentially the same for the U.S. population (Table VI). The data in Table V can be interpreted as follows: Under the child/parent rule at an incidence rate of 1 per 1000 (0.001), the estimator based on the child/parent rule has a smaller MSE than the estimator using a conventional rule for a sample of fewer than 6436 households. If one's sample size requirements necessitate a sample larger than 6436 households, the conventional rule is more efficient.

Table V. Estimated Sample Size for Which the MSE of the Conventional Estimate Equals the MSE of the Multiplicity Estimator at an Incidence Rate of 0.001 for Selected Counting Rules: MSA

	MSE intersection sample size[a]				
Victim subdomain	Sibling	Child, parent	Sibling, child, parent	Friend	Sibling, child, parent, friend
Total sample[b]	1,047	6,436	1,762	2,591	1,758
Crime type					
Burglary	2,063	4,820	2,308	2,539	1,864
Robbery	—[c]	—	1,437	—	2,189
Assault	—	∞	∞	—	19,795
Sex					
Male	1,937	8,340	2,791	8,585	3,185
Female	2,269	23,918	4,439	3,829	3,774
Race					
White	1,180	3,650	1,498	2,054	1,386
Nonwhite	—	—	∞	—	∞
Age					
<35	1,785	∞	4,809	1,857	2,904
≥35	2,629	4,336	2,854	13,542	4,100

[a]The MSE intersection sample size is the estimated sample size where the MSE of the conventional estimator equals the MSE of a specific multiplicity estimator.
[b]$N = 383$.
[c]Fewer than 10 respondents.

Table VI. Estimated Sample Size for Which the MSE of the Conventional Estimate Equals the MSE of the Multiplicity Estimator at an Incidence Rate of 0.001 for Selected Counting Rules: United States

Victim subdomain	MSE intersection sample size[a]				
	Sibling	Child, parent	Sibling, child, parent	Friend	Sibling, child, parent, friend
Total sample[b]	1,020	6,265	1,715	2,521	1,711
Crime type					
Burglary	2,008	4,692	2,246	2,471	1,814
Robbery	—[c]	—	1,399	—	2,131
Assault	—	∞	∞	—	19,268
Sex					
Male	1,876	8,081	2,704	8,319	3,086
Female	2,218	23,378	4,338	3,742	3,688
Race					
White	1,268	3,922	1,609	2,207	1,489
Nonwhite	—	—	∞	—	∞
Age					
<35	1,765	∞	4,756	1,836	2,872
≥35	2,529	4,171	2,746	13,028	3,944

[a]The MSE intersection sample size is the estimated sample size where the MSE of the conventional estimator equals the MSE of a specific multiplicity estimator.
[b]$N = 383$.
[c]Fewer than 10 respondents.

The data in Table V indicate that in about three-quarters of the cells, the conventional rule is the estimator of choice. The reason is that although these sample sizes seem large, they would not yield many respondents who have been victims of a crime. For example, returning to the child/parent rule for the total sample, if we assume two adults per household, an incidence rate of 0.001, and an average of five eligible network households to report each crime victim, a sample of 6436 households would yield about 64 victims if there was perfect reporting. Because the event of victimization is so rare in this example, the resulting sample size would not provide an acceptable sampling error for most analyses and estimation tasks. In general, for a condition this rare based on our results, it is unlikely that network sampling would be the preferred method.

The table cells with infinity indicate that the multiplicity rule is always more efficient than the conventional rule. This occurs for the nonwhite, under-age 35, and assault subdomains. Among the multiplicity rules, the child-parent rule is again the most efficient, but the pattern is not as

dominant as before because a number of the other rules in select instances are more efficient than the conventional rule.

5. DISCUSSION AND CONCLUSIONS

A major objective of our research was to determine how well network respondents—relatives and close friends/co-workers of a crime victim—report crime victimizations. Although it may seem reasonable to conclude that network sampling is not the methodology of choice based on the data that we presented, a number of confounding factors should be addressed. It is difficult to determine how the reverse record check design and the telephone method of data collection, taken together, affected our evaluation of network sampling methodology for crime victimizations.

First, some comments about the reverse record check design using network sampling. In a normal network survey, all respondents report about their own activities and about their eligible networks. The only time that it is necessary to obtain a name, address, or telephone number for another individual is when additional or clarifying information is needed from the nominated person. In our research, it was necessary to interview the victim and then obtain from him or her the names and telephone numbers of a randomly selected relative and of a close friend or co-worker. Our victim respondents were reluctant to provide this identifying information for their networks. Very few respondents refused to give the names or ages of their eligible network members prior to our asking the crime victimization questions. However, when we asked for telephone numbers to contact these individuals, 29% of the victim respondents refused to provide the information for their relatives, and 38% of them refused to give out the telephone numbers of their friends. Many were willing to give identifying information for other eligible network respondents but not for the individuals whom we had randomly selected. Thus, the victim respondents were gatekeepers in our design. However, there is no evidence to suggest that gatekeeping would occur in a normal survey using network sampling. In fact, a number of investigators believe that network respondents may be more willing to report sensitive information about target persons than would the target persons themselves (Sirken, 1974; Bradburn and Sudman, 1979). Finally, we should note that our data showed no relationship between whether victims reported the incident and whether they refused to provide information for contacting the selected network member.

A key factor that we failed to take into account in the design of this study was the nature of the victim population. Unless special steps are taken, such as disproportionate stratification, a RCS will produce a sample

spatially distributed proportional to crime events. This means that the sample will be heavily located in high-crime areas. Although such areas are the appropriate focus of many RCS studies, they are less well suited to testing a network design unless recall aids are incorporated. Many crimes such as burglaries, attempted burglaries, domestic assaults, and possibly robberies are not rare events in some of these areas. Hence, their salience and likelihood of being mentioned to network members may be much lower than in other locations. The ability of respondents to recall a particular event may well be much lower if multiple or similar events have occurred in the same time period. Suspicion about the authenticity of the survey may also be higher in these locations. Clearly, much of this is speculation. What does seem clear is that the issues of what events have high salience and of what factors affect transmission of information through a social network are more complex than we anticipated in our design.

A second confounding factor was the method of data collection. We have conducted three reverse record check network surveys on three different topics using two methods of data collection. One was a study of cancer patients using face-to-face interviews (Czaja *et al.*, 1986), the second concerned missing children and used telephone interviews (Sudman *et al.*, 1987), and the third was the current research. The study to estimate the prevalence of cancer and the cost of care had high reporting rates by both the cancer patient and network households. In addition, only 2% of the patient households refused to provide the network information. In the two telephone surveys, the reporting rates were much lower and refusals to identify network members much higher than in the face-to-face interview survey. It may be an oversimplification to attribute the differences solely to the method of data collection; the method is probably only one of a number of factors contributing to these differences. It is not possible to determine whether the differences between these studies are due to the different methods of data collection or the different survey topics. If network sampling is not appropriate with telephone interviewing in victim surveys and requires face-to-face interviewing, this would have a major impact on the relative costs of victim surveys done with network and traditional sampling, negating some of the efficiencies gained through network sampling. Only further research specifically focused on these issues can answer these questions.

Our results notwithstanding, we believe that there is a need for more experimentation with network sampling. With our sample of victim respondents, 48% had no eligible children or parents, 52% had no eligible siblings, and 12% had no eligible close friends or co-workers. Overall, however, only 2% had no eligible network. The methodology has the potential for increasing the efficiency of a sample to identify persons with rare characteristics. The key is matching the topic with the best counting rules and appropriate method of data collection.

Network sampling also has the potential for identifying cases that would be missed in a conventional survey. In this study, 42% of the victim–network pairs reported the target crime, 25% of the pairs did not report the crime, 28% of the victims reported the target crime when the network member did not, and 5% of the network respondents reported the crime when the victim respondent did not. This last percentage compares favorably with the other studies mentioned above, in which the percentages were between 4 and 8%.

We suggest that research in this area be repeated but with some modifications. We suggest a split method design with half of the interviewing conducted by telephone and the other half conducted face-to-face. We also suggest some experimentation with methods to improve reporting. Recently, there have been efforts to investigate in a laboratory setting how respondent reporting is affected by various types of memory probes (Fathi *et al.*, 1984; Loftus *et al.*, 1985; Royston *et al.*, 1986). Between 5 and 10% of the respondents in our study reported a crime but not the target crime. We suspect that forgetting is the major reason for these nonreports. When persons are victims of multiple crimes, it is very likely that the details of the events may be confused or that one of the events may be forgotten. Care must be taken in guiding the respondents through the reference period and assisting them to recall the details.

We assumed that the crimes we investigated should be salient enough to merit mentioning to network members. It was further assumed that, as part of this mention, network members would obtain other information about the event, such as when it occurred, to make them useful reporters. It may well be that there is an interaction between salience and rarity that is crucial to whether target events are mentioned to network members, and even whether they are long retained by the target respondent. For example, an event such as an attempted burglary may be very salient the day or week after it occurs. However, if it is not also a relatively rare occurrence, it may not be mentioned to, or make much of an impression on, network members. If another similar event occurs later, that later event may "erase/replace" the earlier memory. If an event is rare, such as a petty theft from one's yard, but is not particularly salient, it is not likely to be mentioned to network members or long retained by the victim.

This need for both salience and rarity of events may be crucial in getting good network reports, at least when the event occurs at a single point in time, as opposed to something that occurs over some length of time, such as a chronic health condition or being on probation. The crime events that were the subject of our research are indeed rare over the total population; however, they are much less rare in some neighborhoods. Unfortunately, our sample sizes were not large enough to allow analysis by high versus low crime locations.

One might also reasonably speculate that events whose consequences, in some way, span a period of time may be better reported in network surveys than events of shorter duration. For example, a crime that results in long-term physical or psychological treatment of the victim may be better reported than one, such as minor vandalism, that has few, if any, ongoing consequences. The information about a crime with long-term effects has more time to diffuse through the victim's network, increasing the possibility of a network report; and the victim himself is repeatedly reminded of the event and will be more likely to recall it.

An additional problem is the low reporting for assaults. Many of the assault or battery crimes were domestic disturbances or disputes between acquaintances. After reading many of the case narrations, we suspect that the police were called to help control the situation rather than because the victim thought that a crime had been committed. In some cases, it was unclear who the perpetrator was and who the victim was. When dealing with these and other situations where the respondent does not believe a crime has been committed, it is important to query respondents about any situations where they called the police or where the police were summoned but, in the respondent's mind, no crime had been committed.

A number of the problems that we encountered were due to the quality of the sampling frame and the nature of the target population These need to be addressed in future endeavors. First, the quality of the records was so poor that substantial losses of sample occurred. There is no way to know whether these losses were random across the list or how the inclusion of those might have changed the results. Second, a number of the assault victims were difficult to locate. This may have occurred for several reasons, but it is at least partially due to the mobility of this group. Another factor could be demographic shifting. According to the Chamber of Commerce, using Census estimates and other data, the County lost nearly 9% of its population during the early and mid-1980s. One might speculate that this decline may well have contributed to the difficulties that we experienced in locating respondents in the sample areas. It may well be that this mobility is also an indicator that their contacts with relatives, friends, and co-workers are less frequent than for other sample members. Consequently, fewer of their network members may be aware of things that happened to them, including crime victimizations.

The best counting rules in our research were the child/parent rule and the friend rule. The friend rule provides the broadest coverage but also presents a problem in estimation. There is often not reciprocity between friends. Person A may name person B as a close friend; but B, when asked the same question, may not name A. This creates a problem in estimation because individuals come into the sample with unknown probabilities. We tried to overcome this by asking for a set number of friends, three, but it

did not work. Approximately, 47% of the network friends did not mention the crime victim as one of their three closest friends. Our estimator does not take this into consideration. This issue must be addressed in future work before the friend rule can be an acceptable component in network designs.

ACKNOWLEDGMENTS

This research was supported by a grant from the National Institute of Justice (85-IJ-CX-0032). The authors wish to thank Monroe Sirken, Gad Nathan, and Robert Sampson, who provided helpful comments on early drafts of this paper. We gratefully acknowledge the assistance of Cecelia Snowden, who did the programming for the estimators.

REFERENCES

Biderman, A. D., and Lynch, J. P. (1981). Recency bias in data on self-reported victimization. *Proceedings*, Social Statistics Section, American Statistical Association, pp. 31–40.

Bradburn, N. M., and Sudman, S. (1979). *Improving Interview Method and Questionnaire Design*, Jossey-Bass, San Francisco.

Casady, R. J., Nathan, G., and Sirken, M. G. (1985). Alternative dual system network estimators. *Int. Stat. Rev.* 52: 183–197.

Czaja, R., Snowden, C., and Casady, R. (1986). Reporting bias and sampling errors in a survey of a rare population using multiplicity counting rules. *J. Am. Stat. Assoc.* 81: 411–419.

Dodge, R. W. (1983). Using record checks. *Proceedings*, Section on Survey Research Methods, American Statistical Association, pp. 680–684.

Fathi, D., Schooler, J., and Loftus, E. (1984). Moving survey problems into the cognitive psychology laboratory. *Proceedings*, Section on Survey Research Methods, American Statistical Association, pp. 19–21.

Fienberg, S. E. (1978). Victimization and the national crime survey: Problems of design and analysis. In Namboodiri, N. K. (ed.), *Survey Sampling and Measurement*, Academic Press, New York.

Fischer, C. S. (1982). *To Dwell Among Friends: Personal Networks in Town and City*, University of Chicago Press, Chicago.

Flanagan, T. J., and McLeod, M. (eds.) (1983). *Sourcebook of Criminal Justice Statistics—1982*, Bureau of Justice Statistics, U.S. Department of Justice, U.S. Government Printing Office, Washington, D.C.

Garofalo, J. (1977). *Local Victim Surveys: A Review of the Issues*, U.S. Government Printing Office, Washington, D.C.

Groves, R. M., and Kahn, R. L. (1979). *Surveys by Telephone: A National Comparison with Personal Interviews*, Academic Press, New York.

Kalish, C. B. (1974). *Crimes and Victims: A Report on the Dayton-San Jose Pilot Survey of Victimization*, Law Enforcement Assistance Administration, Washington, D.C.

Loftus, E. F., Fienberg, S. E., and Tanur, J. M. (1985). Cognitive psychology meets the national survey. *Am. Psychol.* 40: 175–180.

Lopata, H. Z. (1979). *Women as Widows: Support Systems*, Elsevier, New York.

Miller, P. V., and Groves, R. M. (1985). Matching survey responses to official records: Validity in victimization reporting. *Publ. Opin. Q.* 49: 366–380.

Nathan, G. (1976). An empirical study of response and sampling errors for multiplicity estimates with different counting rules. *J. Am. Stat. Assoc.* 71: 808–815.

Rittenhouse, J. D., and Sirken, M. G. (1980). A note on networks, nominations and multiplicity as contributory to heroin estimation. In Rittenhouse, J. D. (ed.), *Developmental Papers: Attempts to Improve the Measurement of Heroin Use in the National Survey,* National Institute on Drug Abuse, Rockville, Maryland.

Rothbart, G. S., Fine, M., and Sudman, S. (1982). On finding and interviewing the needles in the haystack: The use of multiplicity sampling. *Publ. Opin. Q.* 46: 408–421.

Royston, P., Bercini, D., Sirken, M., and Mingay, D. (1986). Questionnaire Design Research Laboratory. *Proceedings,* Section on Survey Research Methods, American Statistical Association, pp. 703–707.

Sirken, M. G. (1970). Household surveys with multiplicity. *J. Am. Stat. Assoc.* 65: 257–266.

Sirken, M. G. (1974). The counting rule strategy in sample surveys. *Proceedings,* Social Statistics Section, American Statistical Association, pp. 119–123.

Sirken, M. G. (1979). A dual system network estimator. *Proceedings,* Section on Survey Research Methods, American Statistical Association, pp. 340–342.

Sirken, M. G., and Royston, P. (1977). Counting rule bias in household surveys of deaths. *Proceedings,* Social Statistics Section, American Statistical Association, pp. 347–351.

Sirken, M. G., Inderfurth, G. P., Burnham, C. E., and Danchik, K. M. (1975). Household sample surveys of diabetes: Design effects of counting rules. *Proceedings,* Social Statistics Section, American Statistical Association, pp. 659–663.

Skogan, W. G. (ed.) (1976). *Sample Surveys of the Victims of Crime,* Ballinger, Cambridge, Mass.

Skogan, W. G. (1981). *Issues in the Measurement of Victimization,* Bureau of Justice Statistics, U.S. Department of Justice, U.S. Government Printing Office, Washington, D.C.

Skogan, W. G., and Maxfield, M. G. (1981). *Coping with Crime,* Sage, Beverly Hills, Calif.

Sparks, R. F. (1982). *Research on Victims of Crime: Accomplishments, Issues, and New Directions,* National Institute of Mental Health, Bethesda, Md.

Sudman, S., and Freeman, H. E. (1988). The use of network sampling for locating the seriously ill. *Med. Care* 26: 992–999.

Sudman, S., Cowan, C. D., Blair, J., and Khodadadi, K. (1987). *Studying the Incidence of Missing Children by Special Survey Methods: Final Report,* prepared for the Office of Juvenile Justice and Deliquency Prevention, Office of Juvenile Assistance, Research and Statistics, U.S. Department of Justice. Survey Research Laboratory, University of Illinois, Urbana.

Sudman, S., Sirken, M. G., and Cowan, C. D. (1988). Sampling rare and elusive populations. *Science* 240: 991–996.

Tuchfarber, A. J., and Klecka, W. R. (1976). *Random Digit Dialing: Lowering the Cost of Victimization Surveys,* Police Foundation, Washington, D.C.

Turner, A. G. (1972). *San Jose Methods Test of Known Crime Victims,* Statistics Technical Report No. 1, Law Enforcement Assistance Administration, Washington, D.C.

U.S. Bureau of the Census (1982a). *1980 Census of Housing, Vol. 1. Characteristics of Housing Units, Chap. A. General Housing Characteristics, Part 15. Illinois (HC80-1-A15),* U.S. Government Printing Office, Washington, D.C.

U.S. Bureau of the Census (1982b). *1980 Census of Population, Vol. 1. Characteristics of the Population, Chap. B. General Population Characteristics, Part 15. Illinois (PC80-1-B15),* U.S. Government Printing Office, Washington, D.C.

U.S. Bureau of the Census (1983). *1980 Census of Housing, Vol. 1. Characteristics of Housing Units, Chap. B. Detailed Housing Characteristics, Part 15. Illinois (HC80-1-B15),* U.S. Government Printing Office, Washington, D.C.

U.S. Federal Bureau of Investigation (1987). *Uniform Crime Reports for the U.S.,* U.S. Government Printing Office, Washington, D.C.