**Questionnaire Biases of Illicit Drug Use Survey Estimates That
Depend On Reports of Close Friends**

Monroe G. Sirken, National Center for Health Statistics
5535 Belcrest Road, Room 700, Hyattsville, MD 20782

Abstract

Questionnaire bias occurs when survey questionnaires fail to solicit the items of information that are required for unbiased estimation. In this paper, we derive equations for the questionnaire biases of illicit drug use estimates obtained in three household surveys in which drug users are reported by their close friends. We compare the biased and unbiased versions close friends questionnaires, and note that the biased versions of the questionnaires appear far more respondent friendly and less burdensome than the unbiased questionnaire versions. Also, we determine the conditions that eliminate questionnaire biases, and demonstrate that these conditions are satisfied when the configurations of close friendship linkages of drug users have specified symmetric properties.

In the absence of empirical data, we speculate (1) that response error reductions due to using biased instead of unbiased close friends' questionnaires may more than compensate for the questionnaire biases particularly when drug users' close friendship linkages exhibit the kinds of symmetric properties noted above, and (2) that surveys that depend on close friends' reports of drug users may have the potential capability of improving the quality of drug use estimates in conventional surveys that depend on self reports of drug users.

Key words: questionnaire design, friendship counting rules, network sampling, configurations of friendship linkages.

A. Introduction

In conventionally designed household sample surveys of illicit drug use, drug users are enumerated at their de jure residences and respond for themselves. Exhibit A is an example of the self respondent questionnaire (SRQ).

Exhibit A. Unbiased Self Respondent
Questionnire (SRQ)

"Have you used [illicit drug] at any time during the past year?"

Now consider a household sample survey in which drug users do not self report and instead drug users are reported by close friends. Exhibit B is an example of a close friends questionnaire (CFQ*)..

Exhibit B. Unbiased Close Friends
Questionnaire (CFQ*)

"How many of your close friends used [illicit drug] during the past year? (Don't mention their names.) Answer the next question for each close friend. How many besides yourself consider (...) a close friend and know (...) used [illicit drug] during the past year".

Note that the CFQ* does not ask for the names or other identifiers of the reported drug users.

Questionnaires A and B collect the items of information that satisfy the data requirements for unbiased estimation of the number of people that are users of (illicit drug). Unfortunately, both questionnaires are very prone to response errors though for different reasons. The unbiased SRQ asks for very sensitive personal information that drug users can provide but are often reluctant to provide. The unbiased CFQ* asks for relatively insensitive items of information about drug users that close friends often are unable to provide.

We focus on versions of the CFQ* that do not satisfy the data requirements for unbiased estimation of the number of illicit drug users. These biased questionnaires were used in independently designed household surveys that were conducted by the Office of Substance Abuse, the Michigan State Department of Health (Sirken,1975), the National Institute of Drug Addiction (Rittenhouse and Sirken, 1981), and the National Opinion Research Center (Sudman et al.,1977). We derive questionnaire bias formulas for each of these surveys, and show that the directions and magnitudes of the biases depend on the kinds of linkages

that exist between drug users and close friends. Three kinds of linkages between drug users and close friends that minimize the questionnaire biases of the 3 surveys are described in the next section.

B. Symmetries of close friendship networks

The close friendship network of drug user (...) contains the set of one or more people that are close friends of (...) either because they consider (...) their close friend or (...) considers them its close friends. We define three kinds of symmetry in the friendship network of drug user (...).

Regular Symmetry I - parity in the total number of people that (...) views as his close friends and in the total number of people that consider (...) to be their close friend;

Regular Symmetry II - parity in the total number of people that view (...) as their close friend, and in the total number of people that each of them views as his close friends;.

Super Symmetry - satisfies the combined conditions of regular symmetry I and II.

Consider the case in which a total of 3 people A, B, and C .view (...) as their close friend. Then, (...)'s friendship network has regular symmetry I with parity 3 if (...) has a total 3 close friends, not necessarily, A, B and C; (...)'s friendship network has regular symmetry II with parity 3, if A, B and C each have a total of 3 close friends, not necessarily with each other; and (...)'s network has super symmetry with parity 3 if A, B, and C and (...) each have a total of 3 close friends. For example, the conditions of regular symmetry I parity 3 are satisfied if the friendships of A, B and C with (...) are reciprocated by (...); the conditions of regular symmetry II parity 3 are satisfied if A, B, and C have reciprocal friendships with each other; and the conditions of super symmetry parity 3 are satisfied if A, B, C, and (...) have reciprocal friendships with each other.

C. Unbiased estimation

The sampling frame H = { $H_1,...,H_i,...,H_L$ } contains L households, and household H_i (i = 1, ..., L) is the residence of M_i persons ($P_{i1}..P_{ij}..P_{iM_i}$). Let M = $\sum_i M_i$ = total population size. Of the M persons in the population, a subset of N persons { $I_1,...,I_\alpha,...,I_N$ } are drug users. Let the indicator variable $\delta_{\alpha,ij} = 1$ if drug user I_α (α = 1, ..., N) is eligible to be reported as a friend by person ij (i = 1, ..., L) (j = 1, ..., M_j) and $\delta_{\alpha,ij} = 0$ otherwise. Let $s_\alpha = \sum_i \sum_j \delta_{\alpha,ij}$ = the number of people eligible to report I_α (α = 1, ..., N}; it is referred to as the multiplicity of I_α .

Assuming every drug user has at least one eligible friendship, the unbiased estimator of N in a single stage cluster sample of l households is

$$\hat{N} = \frac{L}{l} \sum_{i=1}^{l} \sum_{j=1}^{M_i} \lambda_{ij} \qquad (1)$$

whe re $\lambda_{ij} = \sum_\alpha \delta_{\alpha,ij}/s_\alpha$ = the weighted sum of the number of drug users eligible to be reported by person ij, (i = 1, ...l) (j = 1 , ..., M_i), $s_\alpha \geq 1$ (α = 1,...,N) .The estimator of N that depends on the SRQ is a special case of equation (1) when $s_\alpha = 1$ (α = 1, ..., N). Clearly, $\hat{N}$ is an unbiased estimator of N, $E(\hat{N}) = \sum_\alpha 1/s_\alpha \sum_i \sum_j \delta_{\alpha,ij} = N$.

The data requirements for unbiased estimation of N are satisfied when survey respondents report the number of their close friends that are drug users and the number of other people that also close friends of each reported drug user.

Let the indicator variable $\delta_{\alpha,ij} = \delta^*_{\alpha,ij}$ if person ij considers drug user I_α as his close friend, and let $\delta_{\alpha,ij} = \delta^{**}_{\alpha,ij}$ if drug user I_α considers person ij as his close friend. The unbiased questionnaire version CFQ* shown earlier in Exhibit B satisfies the data requirement for unbiased estimation when $\delta_{\alpha,ij} = \delta^*_{\alpha,ij}$, and the questionnaire version CFQ ** shown below (Exhibit C) satisfies the data requirements for unbiased estimation when $\delta_{\alpha,ij} = \delta^{**}_{\alpha,ij}$.

Exhibit C. Unbiased Close Friends'
Questionnaire (CFQ**)
"How many people that consider you their close friend used [illicit drug] at any time during past year?" (Don't mention their names.)
Please answer the next question, for each [illicit drug] user that considers you a close friend.
How many of (...)'s other close friends, if any, know that (...) used [illicit drug] during the past year?"

The unbiased CFQ* estimator of N is

$$\hat{N}^* = \frac{L}{l} \sum_{i=1}^{L} \sum_{j=1}^{M_i} \sum_{\alpha=1}^{N} \frac{\delta^*_{\alpha,ij}}{s^*_\alpha}$$

$$s^*_\alpha \geq 1 \; (\alpha = 1, ..., N) \qquad\qquad (2)$$

where $s^*_\alpha = \sum_i \sum_j \delta^*_{\alpha,ij}$ = the number of people that view I_α as a close friend. The unbiased CFQ**estimator of N is

$$\hat{N}^{**} = \frac{L}{l} \sum_{i=1}^{l} \sum_{j=1}^{M_i} \sum_{\alpha=1}^{N} \frac{\delta^{**}_{\alpha,ij}}{s^{**}_\alpha} \qquad s^{**}_\alpha \geq 1$$

$(\alpha = 1, ..., N)$ (3)

where $s_{\alpha}^{**} = \sum_{i=1}^{N} \sum_{j=1}^{M_i} \delta_{\alpha,ij}^{**}$ = the number of people

that I_{α} views as his close friends..

From (2) and (3), it is evident that the CFQ* and CFQ** estimators are equivalent, that is $\hat{N}^* = \hat{N}^{**}$, if the friendship networks of N drug users are all regular symmetric I with parities $s_{\alpha}^{*} = s_{\alpha}^{**}$.(α = 1, ..., N).

D. Questionnaire biases

The questionnaires of the NIDA, Mich, and NORC surveys do not satisfy the data requirements for unbiased estimation

Bias of the NIDA estimator

This is essentially the version of the close friends questionnaire used in the NIDA survey.

Exhibit D. The NIDA Questionnaire

"How many of your close friends do you know for sure used (illicit drug) during the past year? (Don't mention names.)

Please answer the next question for each close friend that was an [illicit drug] user during the past year..

How many of (...)'s other close friends know for sure that (...) used heroin during the past year?"

.

The estimator of N that depends on the NIDA questionnaire is

$$\hat{N}_{NIDA} = \frac{L}{l} \sum_{i=1}^{l} \sum_{j=1}^{M_i} \sum_{\alpha=1}^{N} \frac{\delta_{\alpha,ij}^{*}}{s_{\alpha}^{**}} \quad .$$

$s_{\alpha}^{**} \geq 1$ (α = 1, ..., N).   (4)

The bias of $\hat{N}_{NIDA}$ is

$$Bias_{NIDA} = E(\hat{N}_{NIDA}) - N = \sum_{\alpha=1}^{N} [(\sum_{i=1}^{L} \sum_{j=1}^{M_i} \frac{\delta_{\alpha,ij}^{*}}{s_{\alpha}^{**}}) - 1]$$

$$= \sum_{\alpha=1}^{N} (\frac{s_{\alpha}^{*}}{s_{\alpha}^{**}} - 1). \quad\quad (5)$$

The $Bias_{NIDA} = 0$ if the close friendship networks of N drug users are regular symmetric I with parities $s_{\alpha}^{*} = s_{\alpha}^{**}$ (α = 1, ..., N). The bias is positive if $s_{\alpha}^{*} > s_{\alpha}^{**}$ and negative if $s_{\alpha}^{*} < s_{\alpha}^{**}$ (α = 1, ... , N). Otherwise, the bias may be either positive or negative.

Bias of the Michigan State estimator

The Michigan State close friends questionnaire contains a single question.

Exhibit E. Michigan State questionnaire

"What fraction of your close friends used [illicit drug]

during the past year?"

The estimator of N that depends on the Michigan State questionnaire is

$$\hat{N}_{Mich} = \frac{L}{l} \sum_{i=1}^{l} \sum_{\alpha=1}^{N} \sum_{j=1}^{M_i} \frac{\delta_{\alpha,ij}^{*}}{S_{ij}(\alpha)}$$

$S_{ij}(\alpha) \geq 1$ (6)

where $\sum_{\alpha} \delta_{\alpha,ij}/S_{ij}(\alpha)$ = the fraction of the close friends of person i j (i = 1, ..., L) (j = 1, .., M_i) that are drug users. The bias of $\hat{N}_{Mich}$ is

$$Bias_{Mich} = E(\hat{N}_{Mich}) - N$$
$$= \sum_{\alpha=1}^{N} \sum_{i=1}^{L} \sum_{j=1}^{M} (\frac{1}{S_{ij}(\alpha)} - \frac{1}{s_{\alpha}^{*}}) \delta_{\alpha,ij}^{*}$$

$$S_{ij}(\alpha) \geq 1 \ \ s_{\alpha}^{*} \geq 1 \quad (7)$$

The Michigan State survey estimator of N is biased unless the close friendship networks of the N drug users are regular symmetric II with parities $s_{\alpha}^{*} = S_{ij}(\alpha)$ (j = 1, ... , M_i) (i = 1, ..., L) (α = 1, ..., N). The bias of the estimator is positive if $s_{\alpha}^{*} > S_{ij}(\alpha)$, and negative if $s_{\alpha}^{*} < S_{ij}(\alpha)$ (α = 1, ...N). Otherwise the bias may be positive or negative.

Bias of the NORC estimator

This NORC close friends questionnaire contains a single question.

Exhibit F. The NORC questionnaire

"Think of your three closest friends." (Don't mention their names)

As far as you know how many of them used [illicit drug] during the past year?"

The NORC questionnaire assumes that every person has at least 3 close friends.

The NORC estimator of N is a special case of the Michigan State estimator when $S_{ij}(\alpha) = 3$ (α = 1, ..., N) (i = 1, ..., L) (j = 1, ..., M_i). Substituting $S_{ij}(\alpha) = 3$ in equation 6 and simplifying, the NORC estimator of N is   .

$$\hat{N}_{NORC} = \frac{L}{l} \sum_{\alpha=1}^{N} \frac{s_{\alpha|3}^{*}}{3} \quad\quad (8)$$

where $s_{\alpha|3}^{*}$ = the number of people that view I_{α} (α = 1, ..., N) as one of their 3 closest friends.

The bias of the NORC estimator is

$$Bias_{NORC} = E(\hat{N}_{NORC}) - N = \sum_{\alpha=1}^{N} (\frac{s_{\alpha|3}^*}{3} - 1). \quad (9)$$

The NORC estimator is unbiased if $\sum_{\alpha} s_{\alpha|3}^* = 3N$. The unbiasedness condition is satisfied if the friendship networks of N drug users are all regular symmetric II with parity $s_{\alpha}^* = s_{\alpha|C}^* = S_{ij}(\alpha) = 3$ (i = 1, 2, ..., L) (j = 1, 2, ..., M_i) (α = 1, 2, ..., N).

E. Summary and concluding remarks

.This paper deals with questionnaire biases of estimators of illicit drug use that depend entirely on the counts of drug users reported by close friends in household sample surveys. As defined in this paper, questionnaire bias represents the failure of survey questionnaires to solicit the essential information items as required for unbiased estimation.
.

First, we presented unbiased versions of questionnaires that depend on self reports of drug users (Exhibit A), and that depend the number of drug users reported by their close friends (Exhibits B and C). Though these questionnaires collect the items of information essential for unbiased estimation, they are liable to very large response biases. The self reports questionnaire (Exhibit A) asks for very sensitive kinds of information that drug users are often reluctant to provide, and the close friends questionnaires (Exhibits B and C) ask for relatively insensitive kinds of information that close friends are often unable to provide.

Next, we presented the three biased versions of the close friends questionnaires (Exhibits D, E and F) that were used in drug use surveys that were conducted by the former National Institute of Drug Addiction (NIDA), Michigan State Health Department (Mich) and National Opinion Research Center (NORC). The biased versions of the questionnaire are far more respondent friendly and far less burdensome than unbiased questionnaire versions.

Then, we derived formulas of the questionnaire biases for the estimators of drug use that depend on NIDA, Mich and NORC questionnaires. The direction and magnitude of the questionnaire biases depend on statistical properties of close friendship linkages of drug users. We described 3 kinds of symmetric properties of close friendship linkages of drug users that eliminate questionnaire biases.

Drug use surveys that depend on reports of close friends have the potential capability of improving the quality of conventional drug use estimates that depend on self reports of drug users particularly if the close friendship linkages of illicit drug users have the kinds of symmetric properties that are described in this report. Though the NIDA, Mich and NORC surveys do not provide information about friendship linkages of drug users, the findings of those surveys are encouraging. They indicate that drug use estimates that depend on close friends' reports are substantially larger and more in line with expert opinion than estimates that depend on the self reports of drug users. Also, sampling errors are likely to be smaller, perhaps substantially smaller, because close friends reports of drug users yield larger numbers of drug users than self reports of drug users.(Sirken & Nathan, 1988).

References

Rittenhoue, Joan D. & Sirken, Monroe G. (1981). A note on networks, nominations, and multiplicity as contributory to heroin estimation. In *Developmental Papers: Attempts to Improve the Measurement 0f Heroin Use in the National Survey*. National Institute of Drug Addiction. Rockville, Maryland. pgs 49 - 56.

Sirken, Monroe G. (!975). Evaluation and critique of household surveys of substance use. In *Alcohol and Other Drug Abuse in the State of Michigan.* Michigan State Department of Health. Chapter IV pgs 1 - 36..

Sirken, Monroe G., & Nathan,Gad (1988). Hybrid network estimators. *Proceedings of the American Statistical Association: Survey Research Methods Section*. Pgs. 173-78.

Sudman, S., Blair, E., Bradburn, N., & Stocking, C. (1977). Estimates of threatening behavior based on reports of friends..*Public Opinion Quarterly*, 41, 261-4 .