NONCONTACTS: EFFECT ON SURVEY ESTIMATES

Sara L. Thran and Richard J. Willke, American Medical Association
Sara L. Thran, 515 N. State St., Chicago, IL 60610

In many surveys a subset of the sample is nonlocatable. In surveys of physicians, this problem is not as great as in household surveys, but it cannot be ignored. This study examines the demographic characteristics of physicians who cannot be located by telephone and compares them to respondents, nonrespondents, and ineligibles. The potential effects on survey estimates and implications for change in the sample survey design are considered.

Description of the Socioeconomic Monitoring System

The American Medical Association's Socioeconomic Monitoring System (SMS) is a series of semi-annual telephone surveys of non-federal patient care physicians (excluding resident physicians). The spring survey collects data from approximately 4,000 physicians through an interview averaging 25 minutes in length. The autumn survey collects data from approximately 2,800 respondents through a 16-minute interview.

The sample for each survey is selected from the AMA Physician Masterfile, an enumeration of all physicians in the U.S. The sample design is a stratified random sample with the strata defined by specialty and geographic region. Since 1982, Mathematica Policy Research Inc. (MPR) has conducted all SMS surveys under contract with the AMA. Each survey includes reinterviews with physicians who were initially interviewed a year earlier, as well as interviews with physicians selected for the first time.

The sample files sent to MPR contain current AMA Physician Masterfile information on addresses and telephone numbers. The information on telephone numbers is derived from the AMA Physicians' Professional Activity (PPA) survey, which as of 1987 has been sent to approximately one-quarter of the active physician population each year. In addition, triggered mailings are employed whenever AMA is notified of a change in a physician's status. The PPA survey only has included a question on the respondent's telephone number since 1985. In 1988, telephone numbers were available for 53.2% of the SMS sample from PPA records; in 1989, 76.8% of the sample had telephone numbers from PPA records.

MPR and AMA's Department of Survey Design and Analysis attempt to find telephone numbers for those physicians who have not responded to the PPA as well as for those whose recorded telephone number is incorrect. Telephone look-up involves use of some or all of the following sources: directory assistance, specialty societies, state medical societies, AMA's Group Practice book, AHA directory, and state licensing boards. After checking all of these sources, in 1988 6.1 percent of the sample was nonlocatable and in 1989 4.9 percent was nonlocatable.

Survey response rates have averaged around 70% for the past few years. Even with relatively high response rates, systematic differences between respondents and the remainder of the eligible population may bias survey estimates. Previous studies (Olson, Thran, and Strouse; Thran, Marder, and Willke) have investigated characteristics of SMS survey nonrespondents and pointed to the need for adjustments for nonresponse

bias. Weighting, imputation, and regression techniques were applied to SMS data; for the variables analyzed, the results of all three corrections for nonresponse were similar. It was determined that a weighting strategy would be simplest to implement. The weights were derived by dividing the population and survey respondents into 200 cells defined by specialty (10 categories), years since graduation (5 categories), AMA membership status (2 categories), and board certification status (2 categories). Unit response weights were constructed as the ratio of the proportion of physicians in the population to the proportion of respondents in each cell. Second, an eligibility correction was employed. The eligibility correction divides the subset of the population for which eligibility is known into 40 cells (according to years in practice, AMA membership, sex, and board certification) and calculates the proportion of physicians in each cell who are eligible for the survey (i.e., are nonfederal patient care physicians not in residency training). This defines the eligibility weight. The overall weight applied for a given respondent is the product of a unit response weight and an eligibility weight.

Consideration was also given to developing weights to correct for item nonresponse. However, we found that results weighted to adjust for item nonresponse were quite similar to those adjusted for survey nonresponse. Given the difficulty of creating weights to adjust for nonresponse to each item, we chose to simply use unit weights.

Thus, all possible response statuses have been examined except noncontacts. This study analyzes noncontacts in the 1988 and 1989 spring SMS surveys. The 1988 and 1989 spring surveys were each conducted from mid-March through July.

Analysis

The sample can be divided into four response categories: respondent, nonrespondent, ineligible and noncontact. Respondents are those individuals who are reached by telephone and who complete all or part of the interview. Nonrespondents are reached by telephone but refuse to be interviewed. Ineligibles are also reached by telephone but are found to be ineligible for the interview because the individuals are resident physicians, spend less than 20 hours per week in patient care, are federal employees, are deceased, are retired, are temporarily not practicing medicine, are not practicing in the U.S., or are not physicians. Noncontacts are those individuals for whom a telephone number cannot be found; thus their eligibility and response statuses are unknown.

First, we compare the demographic and practice characteristics (obtained from the Masterfile) of noncontacts, ineligibles, survey respondents, and nonrespondents in a univariate context. Next, we examine the probability of being a noncontact using logistic regression. Then, we conduct a least squares regression analysis on net income, total hours in medical activities and number of patient visits per week. Finally, we obtain predicted values of these variables for each response category.

In addition, we are able to infer whether the physician moved during the year just before the survey, and relate such movement to locatability at the time of the survey as well as to other practice characteristics. This inference is based on whether the zipcode of the physician's preferred mailing address changed between the two "year-end" Masterfiles immediately

preceding the survey. For example, for the 1988 survey, zipcodes from the 1986 and 1987 year-end Masterfiles were compared. Changes in zipcodes will usually, but not always, indicate a change in practice site, and so should be viewed with some caution. The year-end Masterfiles are the AMA's best estimate of the actual activity status, address, etc., of each physician as of the end of the calendar year, and incorporate information received for nearly two years subsequent to the end of the year if it is known to apply retroactively. Thus, a recent move by a physician may not be known at the time of the survey but may be captured by the year-end file.

Results

We examined frequency distributions of respondents, nonrespondents, noncontacts and ineligibles by AMA membership status, place of graduation, board certification status, major professional activity, sex, practice location, specialty, census division, age, and a proxy for whether the physician had moved in the year prior to the survey. The four subsets of the sample are significantly different with respect to each characteristic examined. As expected, a much higher proportion of noncontacts (76.5%) are nonmembers of the AMA than either nonrespondents or ineligibles. Noncontacts are more likely to be foreign medical graduates than either nonrespondents or ineligibles. Noncontacts are similar to ineligibles in terms of their board certification status; both groups include a large proportion of physicians who are not certified. Considerably larger proportions of noncontacts are hospital-based, female, and practicing in large metropolitan areas than either nonrespondents or ineligibles. In addition, noncontacts are less likely to be

general surgeons than either nonrespondents or ineligibles. Nonrespondents, noncontacts, and ineligibles have markedly different age distributions; whereas most nonrespondents are between 36 and 55 years of age, most ineligibles are over 55 and noncontacts tend to be younger. The regional distributions of nonrespondents and ineligibles are similar, while noncontacts are more concentrated in the Middle Atlantic and Pacific census divisions. Finally, noncontacts are about twice as likely to have moved in the previous year than survey respondents.

Table 1 presents logistic regression results on the probability of being a noncontact, using the same set of explanatory variables as in the frequency distributions and a dummy variable for survey year. The reference category is male AMA non-member, U.S. medical graduate, non-certified, office-based general/family practitioners practicing in large metropolitan areas in New England in the 1989 SMS sample who have not moved. Even in a multivariate context, most characteristics identified in Table 1 are significantly related to the probability of being a noncontact.

Table 2 presents the regression equations for net income, hours worked and patient visits per week using the same set of demographic and practice characteristics as predictors. It is interesting to compare the significant determinants of locatability seen in Table 2 with the significant determinants of the practice characteristics seen here. To the extent that they coincide, estimates of population characteristics will be biased if noncontacts are not considered. For example, we see that males, AMA members, older physicians and those who did not move in the previous year are both more likely to be contacted and to have higher net incomes. In this

case we would expect non-contacts to have reported relatively low net incomes if they had responded to our surveys, thereby reducing average income estimates.

Based on the known characteristics of item nonrespondents, survey nonrespondents, ineligibles, and noncontacts we are able to predict their responses (using the coefficients from the regression equations and ignoring the potential for selection bias). Mean predicted values by response status are presented in Table 3. Average annual income, hours per week in medical and administrative activities, and total patient visits per week are considerably lower for noncontacts and ineligibles, while predicted values for survey nonrespondents are quite similar to those of item respondents. Item nonrespondents have higher predicted income than any other group.

Table 4 presents estimated mean income, hours worked and patient visits per week weighted, unweighted, and including different groups of response categories. If adjustments were included for locatability as well as survey nonresponse (5th row in Table 4), estimates of overall means would not be drastically different than with the current weighting strategy; in 1988, the average net income would be approximately $1000 higher and in 1989 it would be about $1000 lower than the values reported using the current weighting strategy.

Conclusion

The current weighting strategy adjusts for survey nonresponse and corrects for eligibility. Further attention will be given to changing the weighting methodology to adjust for locatability; however, it may prove to be unnecessary given the increasing telephone coverage of the sample. Over 84% of the 1990 spring SMS sample had telephone numbers available from PPA records, which implies that the proportion of noncontacts may be lower than the previous two years, depending upon the accuracy of PPA telephone information.

Our access to AMA's Physician Masterfile gives us a somewhat unique advantage in examining characteristics of individuals who do not respond to the survey. We have examined characteristics of survey nonrespondents and found that their predicted income is not very different from survey respondents. However, we find that predicted income of noncontacts is quite different from that of survey respondents. In surveys with a higher noncontact rate, this would lead to serious bias in survey estimates.

Table 1
Logistic Regression Results
Probability of Being a Noncontact

Variable	Beta
Intercept	2.2162***
AMA Member	-0.9993***
Foreign Medical Graduate	0.4131***
Board Certified	-0.6112***
Hospital-based	0.4559***
Year88	0.2421**
Internal Medicine	0.1048
General Surgery	-0.3531*
Pediatrics	0.0843
Obstetrics/Gynecology	-0.4318
Radiology	0.5813**
Psychiatry	-0.1524
Anesthesiology	0.6629***
Pathology	0.7402**
Other Specialties	0.6360***
Middle Atlantic	0.3460
E.N. Central	0.3666
W.N. Central	0.3241
S. Atlantic	-0.0290
E.S. Central	0.1743
W.S. Central	-0.5941
Mountain	-0.1596
Pacific	0.7938***
Rural	-0.5326***
Metropolitan < 1,000,000	-0.5431***
Female	0.6998***
Age	-0.1770***
Age squared/100	0.1647***
Moved Last Year	0.6411***

*,**,*** $p < 0.05$, 0.01, and 0.001 respectively.

Table 2
Regression Parameter Estimates

Variable	Annual Net Income Last Year (In Thousands)	Total Hours Per Week	Total Patient Visits Per Week[a]
Intercept	1.86	59.12***	117.30***
Internal Medicine	27.54***	2.33***	-32.35***
General Surgery	85.64***	-1.06	-45.40***
Pediatrics	-2.07	0.06	-4.42
Obstetrics/Gynecology	72.55***	4.74***	-26.74***
Radiology	71.33***	-2.10*	--
Psychiatry	14.22**	-7.96***	--
Anesthesiology	85.74***	-0.36	--
Pathology	29.88***	-7.29***	--
Other Specialties	30.72***	-5.79***	-28.67***
Middle Atlantic	9.94	0.65	8.66*
E.N. Central	16.42**	2.27**	13.74***
W.N. Central	27.06***	2.96**	39.59***
S. Atlantic	20.92**	2.16*	9.07*
E.S. Central	16.46***	-0.39	12.85***
W.S. Central	7.05	1.76	17.90***
Mountain	-3.07	2.47*	3.35
Pacific	9.24	0.43	-3.89
Female	-29.18***	-5.94***	-21.64***
AMA Member	18.99***	3.95***	11.52***
FMG	-5.98	2.31***	-11.98***
Board Certified	22.44***	-0.18	-1.28
Years of Experience	5.31***	0.04	1.69***
(Years of Experience)^2	-0.115***	-0.007***	-0.049***
Hospital-based	-15.56***	-1.64**	-15.51***
Year88	-9.84***	-0.58	-0.59
Moved Last Year	-26.27***	-0.76	-15.41***
F-value	71.51***	37.15***	39.88***
Adjusted R^2	21.80%	10.08%	11.91%

*,**,*** $p < 0.05$, 0.01, 0.001 respectively.
[a]Excludes radiologists, psychiatrists, anesthesiologists and pathologists.

Table 3
Predicted Values by Response Status

| | Annual Net Income Last Year | | | | Total Hours Per Week | | Total Patient Visits Per Week [a] | |
| | 1988 | | 1989 | | | | | |
	N	Mean	N	Mean	N	Mean	N	Mean
Partial/Complete	4151	$137,273	4111	$146,594	8262	58.64	6576	121.98
Item response	3120	133,996	3151	43,937	7945	58.65	5995	122.21
Item nonresponse	1031	147,185	960	155,316	317	58.29	581	119.56
Survey nonresponse	1687	139,436	1578	148,589	3265	58.84	2747	119.77
Noncontact	413	104,397	328	110,665	741	55.78	555	108.18
Ineligible	546	100,585	648	109,471	1194	53.45	1000	103.83

[a]Excludes radiologists, psychiatrists, anesthesiologists, and pathologists.

Table 4
Estimated Dependent Variable Means

	Annual Net Income Last Year		Total Hours Per Week		Total Patient Visits Per Week[a]	
	1988	1989	1988	1989	1988	1989
Using current weighting strategy	$132,300	$144,700	58.2	58.8	121.1	121.6
Unweighted (item respondents only)	134,000	145,200	58.4	59.0	122.0	123.2
Estimated - including survey nonrespondents	136,000	146,000	58.4	59.1	122.4	123.1
Estimated - including survey nonrespondents and item nonrespondents	137,900	147,700	58.4	59.1	125.9	125.0
Estimated - including survey nonrespondents and noncontacts	133,500	143,600	58.2	58.9	122.0	121.9
Estimated - all response categories except ineligible	135,700	145,600	58.3	58.9	125.3	123.4

[a]Excludes radiologists, psychiatrists, anesthesiologists, and pathologists.