A COMPARISON OF CHARACTERISTICS BETWEEN LATE/DIFFICULT AND NON-LATE/DIFFICULT INTERVIEWS IN THE NATIONAL HEALTH INTERVIEW SURVEY

Pei-Lu Chiu, Howard Riddick, and Ann M. Hardy[1]
National Center for Health Statistics, 6525 Belcrest Rd. Hyattsville, MD 20782

Key Words: Nonresponse, Personal interviews, CAPI, Survey bias

ABSTRACT

Recent studies have shown that "late" or "difficult" interviews comprising the last few percent of the survey interviews are significantly different from the "not-late" or "easy" interviews in their household and person-level characteristics. With tight closeout dates imposed upon surveys such as the National Health Interview Survey (NHIS, an annual national health survey sponsored by the National Center for Health Statistics, Centers for Disease Control and Prevention) and increasing negative feelings toward participation in surveys by the public, "late" and "difficult" cases may potentially become nonrespondents and may subsequently affect the estimates. In the 1998 NHIS, 1,197 interviews out of 38,209 interviewed households were completed after the official interviewer closeout date (15 days beyond the assignment starting date) and identified as "late" interviews. 1,475 interviews were classified "difficult" interviews because they required more than 9 contacts. We found that characteristics of these late/difficult interviews are different from the non-late/difficult cases at both the household and person level. The late/difficult interviews are more likely to be households located in a central city, occupied by a single person, and being rented. These households are less likely to have residents 65 years of age or over with any limitation in activity. We also found that estimates from selected health items are quite different between the late/difficult and non-late/difficult interviews. In this study, excluding these late/difficult interviews resulted in small differences in the estimates of some selected health related items.

INTRODUCTION

This study was initiated by a request of the Interagency Household Survey Nonresponse Subgroup, organized in 1999, to explore the household characteristics of difficult and late interviews among several household surveys. In addition to content differences, these household surveys vary in their modes of administration, length of survey questionnaires, and time in the field for data collection. Thus, the household characteristics of the late or difficult interviews among these household surveys may vary.

In the absence of interviewers persistently visiting non-contact households and making refusal conversions, many of these late or difficult interviews may end up as non-interviews. With the recently increasing survey nonresponse rates in many government surveys, late or difficult cases may bias the survey estimates, and can be a concern if they become nonresponses, and their estimated means may be very different from the rest of the interviewed cases.

The household nonresponse rate in the National Health Interview Survey has been increasing over the past several years. The total Type A nonresponse rate (eligible households with no interviews or insufficient data due to refusal, no one home, language or other problems) has increased more than 6 percentage points from 1995 to 1999. Several hypotheses regarding the increase in NHIS nonresponse rates have been postulated, including the switch from Paper And Pencil Interviewing (PAPI) to Computer Assisted Personal Interviewing (CAPI), increase in the length of the core questionnaire, overall increase in nonresponse seen by many surveys, difficulty in finding people at home, etc. Various efforts have been implemented to increase the response rate, such as extending the close out date in the field office to allow more time for refusal conversion and contacts, increasing the number of visits, and giving incentives to the interviewers. With the increasing reluctance of respondents towards survey participation observed in the last decade and the increasing numbers of respondents that cannot be reached before the survey close out date, it has become an increasing challenge for the NHIS interviewers to maintain high response rates.

Who are the potential nonrespondents that the interviewers should concentrate their effort on? What is the effect of excluding these potential nonresponding cases? Kennickell (1999) found that high-income households tend to be interviewed later in the Survey

[1] Current address for Ann Hardy is Social Sciences, Nursing, Epidemiology, and Methods (SNEM) Integrated Review Group, Center for Scientific Review, NIH 6701 Rockledge Dr., Room 3158, MSC 7770, Bethesda MD 20892

of Consumer Finances. If the "head" of the household is over 65, has an education level less than a high school degree, or is currently unemployed, then the household is more likely to be interviewed earlier. Bates and Creighton (2000) examined the household and person level characteristics of late or difficult interviews in the 1999 National Crime Victimization Survey and 1999 Current Population Survey and suggested that the late cases are somewhat similar to nonrespondents. They also found statistically significant differences in the labor force participation rate, unemployment rate, and several types of crime rates from some of the race/age subgroups if the estimates were re-calculated with "difficult" cases excluded. In the 1996 Medical Expenditure Panel Survey, Cohen et al. (2000) found that excluding nonrespondents can reduce the precision of the estimate, but not substantially. Keeter et al. (2000) compared the result of 91 question items from telephone interviews using "Standard" and "Rigorous" methods, which resulted in 30% and 60.6% response rates respectively, and found few significant differences. When assessing the impact of response rate on the Index of Consumer Sentiment (ICS), Curtin, et al. (2000) found that respondents interviewed with fewer calls differed from those interviewed later requiring more calls. However, they found small differences in the estimates of ICS when respondents requiring more calls were excluded.

In this study, we examine differences between the late/difficult and non-late/difficult interviews with respect to their household and person-level characteristics. If there are differences, we want to explore the predictive factors associated with the late/difficult interviews. We also evaluate the effect of excluding the late/difficult interviews in selected health measures.

METHODOLOGY

Data Source

The National Health Interview Survey, conducted by the National Center for Health Statistics (NCHS) annually since 1957, is the Nation's primary source of general health information on the civilian, non-institutionalized, household population of the United States. The data are collected under a contractual relationship with the U.S. Bureau of the Census. A PAPI questionnaire was used from 1957 to 1996. Beginning in 1997, the survey questionnaire was redesigned and converted to CAPI. The redesigned basic core questionnaire is composed of five major sections. The Household Composition Section collects basic demographic information on the household members through a household respondent. The Family Core Section collects health information on all the family members through a knowledgeable member of the family. The Sample Adult Core Section collects health information on a randomly selected adult in the family through self-reporting. The Sample Child Section collects health information on a randomly selected child in the family through a knowledgeable family member. Finally, the Immunization Section collects detailed shot records from sample children and children 12 to 35 months old. In different data collection years, additional sets of questions (called topical or periodic modules) have been added to the basic core questionnaire. In 1998, three topical modules (Adult Prevention, Child Prevention, and Pregnancy and Smoking) were added to assess progress toward the national Healthy People Objectives for 2000.

Interview Outcome

In 1998, a total of 71,938 households were contacted. Among them, Type C non-interviews (e.g. demolished, condemned households, or households converted to business) constituted 5.2% of the total sample households. About 35.8% of the total sample households resulted in Type B non-interviews, including vacant households (10.4%), households screened out by the sample design feature for the over-sampling of blacks and Hispanics through the Household Composition part of the questionnaire (22.5%), households entirely occupied by military personnel (0.2%) or persons usually living elsewhere (1.2%), and other Type B non-interviews (1.5%). The remaining 42,469 households were eligible for interview. About 8% of the eligible households were Type A non-interviews determined at the field level. Through an in-house file editing process, additional households with insufficient data were rejected as non-acceptable partial interviews and were also classified as Type A non-interviews. This resulted in a total of 38,209 acceptable interviewed households (response rate = 90.0%).

Since the NHIS is used as a sampling frame for other surveys, obtaining demographic information for a household member is essential and considered to be a minimal requirement for acceptable partial interviews at the field level. This is accomplished through the household composition section of the survey. The household composition section also serves to determine the eligibility of the household members to be included in the NHIS. However, the NHIS in-house definition for an acceptable partial interview requires more information than just the household members' demographic and family structure. An NHIS acceptable partially interviewed household is defined as one in which at least one family member in the household has completed up to the "education attainment" question in the middle of the fifth section

of the family core questionnaire (the Family Socio-demographic Section). This means that the acceptable household will have provided information on topics such as general health status, activity limitation, injury, poisoning, access and utilization, health insurance, citizenship, and education attainment for at least one household family. To be classified as a complete interview, the household has to finish the family core, sample adult core, sample child core, and immunization questions for all the families in the household. In this study, we used the in-house definition for complete and acceptable partial interviews.

Definition of Late Interviews

The NHIS interviewers were given 2 weeks and 2 days from the first day of the assignment week to complete the interview. For example, if we designate the assignment week as Monday, January 19 to Sunday, January 25, then the close out date for that week of assignments would be Tuesday, February 3 for the interviewers. This represents a total of 16 workdays because case transmission can occur before midnight on February 3 (15 days beyond the starting date). Those cases requiring additional time to complete or additional effort for refusal conversion by experienced interviewers were allowed to stay in the field after 15 days beyond the starting date, and these cases were considered to be "late" interviews in this study. There were a total of 1,197 households out of the 38,209 acceptable interviewed households in this category. Assuming that the close out date was strictly followed, all of these late interviews would have become Type A non-interviews. Although the interviewers were encouraged to make the first household visit during the first week of the case assignment, we found that nearly 17% of the late interviews had only 1 reported personal visit by the interviewers. In 1998, over 54% of the acceptable interviews were completed in the first week of the assignment.

Definition of Difficult Interviews

The definition of a "difficult" interview is based on the number of personal visits required by the interviewer to complete or close the interview. We plotted the interviewer self-reported[2] "number of personal visits" and chose a visit number that is close to the inflection point as the cut-off. In this study, when a household interview required 10 or more visits by the interviewer, it was considered a "difficult" interview. There were 1,475 households identified as difficult households by the above definition, about 4% of the total acceptable interviewed cases. A caveat to the "number of personal visits" is that there is no way for us to differentiate between a visit resulting in an actual contact with a person and a visit in which the interviewer simply knocked on the door. Our instrument did not keep a record of the nature of each visit. Therefore, a "difficult" interview can be either an interview requiring repeated visits to complete or an interview in which it was hard to find anyone at home. If the interviewer did not persistently visit the household, the household could become a non-interview or non-acceptable partial interview at the close out date.

Analysis

We examined the characteristics of the late and difficult interviewed households and found that many of the characteristic differences between the late and non-late cases are often similar to those between the difficult and non-difficult cases. Since HIS CAPI did not maintain detailed records on the nature and outcome of each interview visit for each interviewed household, we cannot differentiate which of these late or difficult interviews are potential not-at-homes or soft refusals. Thus, we decided to combine the difficult and the late interviews for the comparison studies and grouped these cases as late/difficult interviews. The remaining interviews, completed before the close out date and requiring less than 10 personal visits, were classified as non-late/difficult interviews. To compare the characteristics of the late/difficult and non-late/difficult households, Chi-square tests were carried out using SUDAAN software and the interim weight (weight with post-stratification adjustment only). To understand the predictive power of various household characteristics on the late/difficult interviews, a logistic regression model was used. We also examined the difference between the original estimates (estimates from all interviewed cases) for several selected health items and the estimates from the re-weighted dataset that has late/difficult cases removed as type A non-interviews. The health estimates from the above two datasets were performed using SUDAAN with appropriate final person weights.

RESULTS

Characteristics of Late/difficult Interviews

Late/difficult interviews were more likely than non-late/difficult ones to be partially completed. Table 1 shows that more than 46% of the late/difficult interviews were partially completed, while only 17.6%

[2] In the 1998 NHIS, the interviewer was asked the following question each time before exiting the CAPI instrument. The interviewer filled in an answer based on their recollection. "How many times have you attempted personal contact (actually visited the address) at this address?"

the non-late/difficult interviews were partial interviews.

Table 1. A Comparison of the Household Characteristics between the Late/Difficult and Non-Late/Difficult Interviews from the 1998 NHIS

Household-level Characteristics	Non-Late/ Difficult	Late/ Difficult
Interview completion status*		
Completed Interviews	82.39%	53.94%
Partially completed Interviews	17.61%	46.06%
Primary mode of interview*		
Personal visits	82.70%	66.39%
Telephone	17.30%	33.61%
Geographic regions*		
Northeast	19.21%	20.91%
Midwest	25.28%	20.31%
South	36.41%	32.65%
West	19.10%	26.12%
Urbanicity*		
Urban area	72.87%	81.89%
Rural area	27.13%	18.11%
MSA / Non-MSA residence*		
In MSA; In Central City	28.89%	38.06%
In MSA; Not in Central City	49.02%	46.70%
Not in MSA	22.09%	15.25%
Availability of phone		
Has phone	95.47%	94.53%
Does not have phone	4.53%	5.47%
Household size*		
One person	26.91%	32.72%
Two persons	33.30%	29.39%
Three persons	16.19%	15.56%
Four persons	14.14%	13.60%
Five or more persons	9.46%	8.73%
Household types*		
One adult, no child under 18	26.91%	32.72%
Multiple adults, no child under 18	37.30%	29.99%
One adult, one or more child(ren) under 18	5.88%	7.99%
Multiple adults, one or more child(ren) under 18	29.91%	29.30%

* The difference of the percent distributions between non-late/difficult and late/difficult interviewed households is significant at the 0.05 level.

The NHIS required interviewers to make the initial contact in person. The follow-up interviews could be conducted by telephone if a personal visit was not possible. The data in Table 1 show that a higher percentage of the late/difficult interviews were conducted primarily by telephone than non-late/difficult interviews (based on interviewer's self reporting of primary mode of interviews). Our data did not provide information on the mode of interview for each section of the survey. We cannot tell how much of the survey information was collected by telephone or which sections of the survey were likely to be conducted through follow up phone calls. For example, when the sample adult was not the same person as the family respondent, was the sample adult section of the survey a candidate to be conducted by telephone if the interviewer was running short of time to pay a personal visit?

We compared the geographic locations and household structure between the late/difficult and non-late/difficult interviewed households. The results are given in Table 1. The percent of late/difficult household interviews is significantly higher in the Western region of the country, in urban areas, in central cities, and with single adult households.

We further explored the demographic and socio-economic characteristics of the late/difficult and the non-late/difficult interviews based on the household members' characteristics; results are exhibited in Table 1 (continued). Higher percentages of the late/difficult interview households had at least one Hispanic or black than the non-late/difficult interviewed households. The percentage of late/difficult household interviews having elderly residents (age 65 and over) is significantly lower than for the non-late/difficult interviewed households. Since higher percentages of the late/difficult interviewed households were reluctant to provide education information, differences in the tendency to be more highly educated cannot be as clearly determined as differences for some of the other household characteristics.

When comparing household income, almost 14% of the late/difficult-interviewed households did not provide any information on their income, which is nearly 6 percentage points higher than for non-late/difficult interviewed households. The item nonresponse of the detailed income question in the NHIS was very high (over 25%). The purpose of using the $20,000 cutoff in this study is to take advantage of the lower item response rate (less than 10% nonresponse) in the NHIS's follow-up income category question (above/below $20,000).

Although more than 5% of the late/difficult interviewed households did not report their house ownership status, the general pattern remains unchanged after excluding the households with missing house ownership data. Thus, higher percentages of late/difficult interviewed households than non-late/difficult households are renters or have other living arrangements.

The data illustrate that less than 17% of the late/difficult household interviews have one or more persons with some kind of limitation in activities. By contrast, more than 27% of the non-late/difficult interview households have household members with activity limitations. Similarly, there is a lower percentage of late/difficult interviewed households

with household members not working due to health problems than of non-late/difficult interviewed households (12.07% vs. 19.45%). We also found that significantly higher percentages of the household respondents in the non-late/difficult interviewed households have activity limitations or are not working due to health problems, and many of them are elderly also. This suggested that the availability of the elderly household member or someone with a functional limitation at the time of the interviewer's visits could contribute to the observed difference.

Table 1 (continued)**. A Comparison of the Household Characteristics between the Late/Difficult and Non-Late/Difficult Interviews from the 1998 NHIS**

Household-level Characteristics	Non-Late/ Difficult	Late/ Difficult
Race / Ethnicity*		
At least one Hispanic person	10.42%	14.95%
At least one non-Hispanic black person	11.68%	13.87%
Other	77.91%	71.18%
Household with person aged 65 years or older*		
Has person age 65+	24.43%	12.22%
No person age 65+	75.57%	87.78%
Highest education attainment among all household members*		
High school and below	38.31%	31.33%
College	46.81%	48.46%
Above College	11.30%	11.06%
Don't know or Refused	3.59%	9.15%
Household income*		
Total HH income < $20,000	25.16%	22.95%
Total HH income >= $20.000	67.12%	63.50%
Don't know or Refused	7.84%	13.55%
House ownership*		
Owned or being bought	66.46%	51.24%
Rented or other arrangement	31.90%	43.32%
Don't know or Refused	1.64%	5.44%
Number of persons limited in activities*		
One or more persons limited in any activity	27.71%	16.79%
No one limited in any activity	71.50%	81.37%
Don't know or Refused	0.51%	1.84%
Number of persons 18 years old or older who are not working due to health problem*		
One or more persons	19.45%	12.07%
No one	79.62%	85.93%
Don't know or Refused	0.93%	2.00%

* The difference of the percent distributions between non-late/difficult and late/difficult interviewed households is significant at the 0.05 level.

Model Predicting the Likelihood of a Late/difficult Interviewed Household

The frequency distributions presented in Table 1 do not provide the information necessary to determine how these influential variables are related to one another and the relative importance of each attribute. We fitted a series of multivariate logistic regression models predicting the late/difficult interviews to better understand the individual and combined effect of these variables and any existing covariation between the household characteristics.

Table 2. Coefficients and Odds Ratios of Logistic Model Predicting a Late/difficult Interview

	β	Odds ratio
Constant	- 3.09	
Household types		
One adult, no children	0.32*	1.37*
Multiple adults, no children	0.01	1.09
One adult, 1+ child(ren)	0.17	1.19
Multiple adults, 1+ child(ren)	0	1
Geographical region		
Northeast	0.17	1.19
West	0.43*	1.54*
South	0.07	1.07
Midwest	0	1
MSA/Non-MSA residence		
In MSA; in central city	0.37*	1.45*
In MSA; Not in central city	0.15	1.15
Not in MSA	0	1
Highest education attainment among all household members		
High school and below	- 0.12*	0.88*
Above college	0.03	1.03
Some college or college graduates	0	1
House ownership		
Renting or other arrangement	0.36*	1.43*
Own or buying the residence	0	1
Household with person aged 65 years or older		
Has person aged 65 or older	- 0.83*	0.44*
No person aged 65 or older	0	1
Household with person limited in activity		
Has person limited in activity	- 0.44*	0.64*
No person limited in activity	0	1

* Significant at level .05

We used backward elimination to select the best model in predicting the late/difficult interviews

(late/difficult = 1, non-late/difficult = 0). The household characteristics included in the original model are race/ethnicity, household type, region, MSA, urbanicity, highest education of a household member, household income, house ownership, household with person aged 65 and older, and household with person with activity limitation. We found that the two categories of household income (below $20,000 versus $20,000 or above), urban/rural, and race/ethnicity (Hispanic and non-Hispanic black versus white and other race) are not significant predictors. After eliminating these three predictors sequentially, all the rest of the household characteristics were significant in predicting the late/difficult interviews ($p < 0.05$).

Table 2 provides results of the final household model from the multivariate logistic regression. Single-person households are more likely to have late/difficult interviews compared to households with adults and children. The odds of a household becoming a late/difficult interview are more than 50% higher in the Western region than in the Midwest. Households in the central city are more likely to result in late/difficult interviews than in rural areas. Families renting their households are more likely to be late/difficult interviews than homeowners. However, if the education level of the household is lower (high school or below), the likelihood of the household becoming a late/difficult interview is lower than for households including an individual with college level education. Two strong negative predictors for the late/difficult interviews are households with seniors and households with people that have activity limitations.

Health Estimates With and Without Late/difficult Interviews

Assuming that the late and difficult interviews could become non-interviews if additional efforts were not made, how much will their non-inclusion affect the estimates?

As shown in Table 1, the late/difficult interview households are quite different from the rest of the households in their race/ethnicity composition, geographic location, socio-economic status, age profile, and health status. Table 3 provides a comparison between the estimates of 7 health items in the late/difficult and non-late/difficult interviews.

A general observation is that the late/difficult interviews have higher item nonresponse. There are significant differences for all the health items examined. Among them, the percentage of people with health insurance coverage in the late/difficult interviews is nearly 5 percentage points lower than for the non-late/difficult interviews. Large differences are also observed in the health items "any limitation of activity" and "Medicare coverage." Will these differences of late/difficult interviews significantly bias the general health estimates if they become non-interviews? We compared the estimates of these 7 personal health items between two datasets: one dataset that contained late/difficult interviews, and another dataset that excluded the late/difficult interviews. In the dataset without the late/difficult interviews, the household and person weights were re-calculated by treating the late/difficult interviews as Type A non-interviews. The comparative results of the selected personal health measures are given in Table 4.

Table 3. A Comparison of Selected Estimates of Health Items between Late/Difficult and Non-Late/Difficult Interviews from the 1998 NHIS

Person-level Health Items	Non-Late/ Difficult	Late/ Difficult
Has any limitation of activity*		
Limited in any way	12.95%	8.23%
Not limited in any way	87.05%	91.77%
Reported health status*		
Excellent	38.09%	39.33%
Very good	30.49%	33.05%
Good	22.15%	19.50%
Fair	6.60%	4.81%
Poor	2.29%	1.55%
Don't know/Refused/Not ascertained	0.47%	1.76%
Saw health professional in office in the past 2 weeks*		
Yes	14.72%	11.86%
No	84.56%	85.75%
Don't know/Refused/Not ascertained	0.72%	2.40%
Has health insurance coverage*		
Yes	84.63%	79.84%
No	14.35%	17.17%
Don't know/Refused/Not ascertained	1.02%	2.99%
Delayed medical care due to cost*		
Yes	6.52%	5.79%
No	92.96%	92.14%
Don't know/Refused/Not ascertained	0.52%	2.08%
Has Medicare coverage*		
Yes	12.66%	6.63%
No	86.32%	90.38%
Don't know/Refused/Not ascertained	1.02%	2.99%
Has Medicaid coverage*		
Yes	8.12%	7.81%
No	90.86%	89.20%
Don't know/Refused/Not ascertained	1.02%	2.99%

* The difference of the percent distributions between non-late/difficult and late/difficult interviewed households is significant at the 0.05 level.

The differences in the selected health items varied from 0 to 0.15 percentage points. The health measures with larger differences were correlated with those with large differences between the late/difficult and the non-late/difficult interviews. For example, higher percentages of households with people having activity

limitation are in the non-late/difficult interview category; thus, the estimate for households having people with any limitation of activity is slightly higher in the dataset with late/difficult interviews excluded. However, for Medicare coverage, there is no difference in the final estimates between all cases and the cases with late/difficult interviews excluded. Medicare coverage is highly correlated to age. The lower number of elders in the late/difficult interviews did not affect the final estimates.

Table 4. Estimates of Personal Health Related Items for Datasets with All Interviewed Cases and Datasets Excluding Late/Difficult Interviews

Person-level Health Items	All Cases*	Exclude Late/Difficult*
Has any limitation of activity	12.65% (0.15)	12.80 % (0.15)
Reported health status		
Excellent	38.17% (0.30)	38.20% (0.31)
Very good	30.57% (0.24)	30.47% (0.25)
Good	21.99% (0.22)	22.07% (0.23)
Fair	6.48% (0.10)	6.54% (0.11)
Poor	2.25% (0.06)	2.25% (0.06)
Saw health professional in office in the past 2 weeks	14.54% (0.14)	14.66% (0.14)
Has health insurance coverage	84.33% (0.23)	84.44% (0.24)
Delayed medical care due to cost	6.47% (0.13)	6.57% (0.14)
Has Medicare coverage	12.27% (0.19)	12.29% (0.20)
Has Medicaid coverage	8.10% (0.19)	8.19% (0.20)

* Percent with characteristic, with standard error in parentheses

CONCLUSIONS AND DISCUSSION

To answer the question of whether there are differences between the late/difficult and non-late/difficult interviews, our data illustrate that there are definite differences. First of all, the data collected from these late/difficult interviews are of a poorer quality, as evidenced by the fact that over 46% of the late/difficult interviews were only partially completed. The item nonresponse is also higher in the late/difficult interviews. NHIS is a face-to-face personal interview survey, yet nearly one third of the late/difficult interviews were conducted primarily by telephone. Second, the late/difficult interviews are more likely to occur in the West than in the Midwest and are also more likely to occur in central cities. Third, the demographic, socio-economic, and health profiles of the late/difficult interviewed households are different from the non-late/difficult households. This gives rise to our original concern that if the late/difficult interviews become non-interviews, this may affect the estimates. In this study, the late/difficult cases constitute 7% of the total interviewed cases and we observed only small differences. However, the difference may be larger if the comparison is at the subgroup level or among other health items.

How do the characteristics of the late/difficult household interviews compare to the non-interviews in the literature? Nonresponse could come from either refusal or not-at-home, and the household characteristics of these two types of non-interviews may not be the same. Single-person households were reported to have a lower contact rate and cooperation rate (Groves and Cooper, 1998). We found that single person households are more likely to be late/difficult interviews. As for the behavior patterns of racial and ethnic subgroups in the survey, social isolation theory has been adopted as the popular hypothesis. Previous reports on the cooperative behavior of minority groups in surveys are mixed. Hawkins (1975) found fewer nonwhites among the refusals in a face-to-face survey in Detroit. On the other hand, DeMaio (1980) and Smith (1983) found that race has no effect on survey cooperation. Groves and Couper (1998) indicated that a higher cooperative rate among minority groups could be explained away by controlling for their socioeconomic status. We found that the effect of race/ethnicity became insignificant when all other predicting factors were controlled for.

The effect of socioeconomic status on survey participation has been widely studied. The final results have varied from survey to survey. DeMaio (1980) found that low-income households were least likely to refuse survey interviews. We noticed that high item nonresponse for income is associated with late/difficult interviews. Based on the households with income information available, we did not find income as a significant predictor for late/difficult interviews.

Groves and Cooper (1998) found that middle-aged households were less likely to cooperate than young and old households. DeMario (1980) reported that middle-aged and older people cooperate at lower rates than those under 30. Groves and Cooper (1998) suggested that the elderly are more frequently at home due to their low employment rate and reduced mobility; however, their poor health may prevent them from survey participation. We found that households containing seniors and members with activity limitations remain a strong negative influence on predicting the late/difficult interviews when controlling for all other predicting variables. We believe that this is not only because these people are more likely to be home during the day, but also because the topic of health is viewed favorably among the elderly and

people with poorer health. The NHIS is appealing to these respondents, and it is easier to gain cooperation.

The literature has been consistent in documenting the correlation of survey participation with urbanicity. Smith (1983) reported lower response rates in central cities than in other metropolitan areas and rural areas; both the refusals and not-at-homes shared similar patterns. Goyder et al. (1992) concluded that the social disorganization of the inter-city contributes to the low response rate. Groves and Cooper (1998) suggest that population density, higher crime rates, and social disorganization in urban areas have a negative effect on survey cooperation. Our finding that late/difficult interviews are more likely to occur in a central city is consistent with these reports. In general, many of the household characteristics associated with the late/difficult households are similar to those of the non-responding households in the literature.

FUTURE RESEARCH

In this study we define difficult interviews based upon the number of personal contacts required. We cannot differentiate as to whether the high number of contacts is due to not-at-home, difficult to gain cooperation, or requiring a high number of visits for data collection. Similarly, the late interviews were defined based on the duration of the survey in the field; we do not know how much effort the interviewer had to render to gain cooperation. In the future NHIS, we have added questions to assess the degree of cooperation among household respondents based on each interviewer's observation. We would like to use this survey cooperation information to explore the household and person-level characteristics of responding households, and their correlation with late and difficult interviews. We also would like to investigate possible causes of partial interviews. Is break-off caused by asking certain types of questions (e.g., income, citizenship, or immunization questions), excessive length of the survey questionnaire, or the difficulty of getting access to the respondents (e.g. sample adult)? We also would like to look in more depth at the data quality of late and difficult interviews. Is there a higher incidence of item nonresponse or "don't know"/"refuse" responses on certain critical questions, or are these types of responses pervasive throughout the entire set of questions? Finally, we would like to examine the impact of several survey administrative changes in the past few years on the survey participation.

REFERENCES

Bates, N., and Creighton, K. (2000). "The Last Five Percent: What Can We Learn from Late Interviews?" A paper presented at the American Statistical Association Annual Meetings, Indianapolis, Indiana.

Cohen, S.B., and Machlin, S.R., and Branscome, J.M. (2000). "Patterns of Survey Attrition and Reluctant Response in the 1996 Medical Expenditure Panel Survey." *Health Services and Outcomes Research Methodology,* 1 (2): 131-148.

Curtin, R., Presser S., and Singer, E. (2000). "The Effects of Response Rate Change on the Index of Consumer Sentiment." *Public Opinion Quarterly*, 64:413-428.

DeMaio, Y.J. (1980) "Refusal: Who, Where and Why:" *Public Opinion Quarterly*, 44: 223-233.

Goyder, L., Lock, J., and McNair, T. (1992). "Urbanization Effects on Survey Nonresponse: A Test Within and Across Cities." *Quality and Quantity*, 26: 39-48.

Groves, R., and Cooper, M. (1998). "Nonresponse in Household Interview Survey", New York: Wiley.

Hawkins, D.F. (1975) "Estimation of Nonresponse Bias" *Sociological Methods and Research*, 3: 461-488.

Keeter, S., Miller, C., Kohut, A., Groves, R., and Presser, S. (2000). "Consequences of Reducing Nonresponse in a Large National Telephone Survey." *Public Opinion Quarterly*, 64:125-148.

Kennickell, A. (1999). "What Do the "Late" Cases Tell Us: Evidence from the 1998 Survey of Consumer Finances," A paper presented at the International Conference on Survey Nonresponse, Portland, Oregon.

Smith, T.W. (1983). "The Hidden 25 Percent: An Analysis of Nonresponse on the 1980 General Social Survey." *Public Opinion Quarterly*, 47: 386-404.