

The Impact of Providing Incentives to Initial Telephone Survey Refusers on Sample Composition and Data Quality

Doug Currivan

RTI International

This paper was prepared for the American Association of Public Opinion Research annual meetings in Miami, May 12-15, 2005. The author gratefully acknowledges Joanne Pais of RTI for running the analyses presented here; Barbara Bibb of RTI for technical assistance with the data file; and project director Matthew Farrelly of RTI and the New York Department of Public Health for supporting this research. All conclusions and opinions expressed in this paper are solely those of the author.

The Impact of Providing Incentives to Initial Telephone Survey Refusers on Sample Composition and Data Quality

Although much research has examined the impact of standard incentive plans on survey response, relatively few studies have evaluated the impact of offering incentives to initial refusers on the final composition of the sample, the distribution of survey responses, or the quality of survey data. As usually intended, offering an incentive payment based on initial refusal may increase the participation of sample members who would otherwise be difficult to include in the survey. Research also suggests that incentives may persuade sample members who have little interest in the survey topic to participate. While greater participation among less interested sample members may reduce unit nonresponse bias, respondents with little interest in the survey topic might also be less motivated to provide complete and accurate responses than other sample members. Combined, these two factors could result in increased participation from some sample members, but poorer quality data from these respondents. To investigate the impact of offering incentives to initial refusers, this study evaluates an incentive protocol currently being implemented in a statewide telephone survey, the New York Adult Tobacco Survey. In this survey, respondents are not initially offered any incentive to complete the interview, but are then offered \$20 if they initially refuse to participate. Comparisons of key characteristics and survey responses among sample members who received the incentive versus those who did not indicate how this payment influenced the composition of respondents. Comparisons of response patterns to key survey items among paid and unpaid respondents – including analysis of missing data, “don’t know” answers, and other indicators of low respondent effort – indicate whether the incentive affected data quality. The implications of these analyses for similar incentive protocols are discussed.

Introduction

A considerable amount of research over the past several years has examined the impact of various incentive amounts and protocols on survey participation. Key reviews by Church (1993) on the impact of incentives in self-administered mail surveys and Singer, Van Hoewyk, Gebler, Raghunathan, and McGonagle (1999a) on the impact of incentives in interviewer-administered surveys have greatly expanded our knowledge of the effectiveness of incentives under different survey modes conditions. As Singer (2002) points out, this knowledge base focuses mostly on the primary “intended” consequence of using incentives – to use available survey resources to maximize the response rate. Relatively few studies have evaluated the impact of offering incentives to initial refusers on survey participation or other survey outcomes. A recent study by

Singer, Groves, and Corning (1999b) represents one of the few attempts to examine the “unintended” consequences of using incentives, including the sample composition, the quality of responses, concerns about equity, and the development of expectation effects (see also Groves, Singer, Corning, and Bowers, 1999 for a related study).

A primary concern about potential unintended consequences stems mainly, but not exclusively, from the use of differential incentives. Differential incentive refers to a survey protocol where sample members are offered different amounts based on either their characteristics or behaviors. As one example, a survey may offer an incentive only after sample members have refused or otherwise expressed reluctance to participate. Offering different sample members different incentive amounts may make economic and practical sense in achieving study goals, but may also produce unintended consequences such as those Groves, et al (1999) and Singer, et al (1999b) investigated and Singer (2002) reviewed. Unintended consequences include introducing bias in survey estimates by altering sample composition, lowering item response quality among paid respondents, encouraging sample members to refuse until payment is offered, and provoking feelings of inequity among unpaid respondents.

To investigate the intended and unintended consequences of a survey protocol offering incentives to initial refusers, this study evaluates an incentive plan currently being implemented in a statewide telephone survey, the New York Adult Tobacco Survey (NY ATS). In this survey, respondents are not initially offered an incentive payment to complete the interview, but are then offered \$20 if they initially refuse to participate. This paper looks at the impact of using these refusal conversion incentives on the composition of survey respondents, the distribution of their responses, and the quality of responses they provide. Comparisons of key characteristics and survey responses among sample members who received the refusal conversion incentive versus those who did not will indicate how this payment influenced the composition of respondents. Comparisons of response patterns to key survey items among paid and unpaid respondents – including analysis of missing data, “don’t know” answers, and other indicators of respondent effort – will provide some preliminary evidence on whether the incentive payment affected data quality among initial refusers. The implications of the findings for similar incentive schemes are discussed.

Refusal Conversion Incentives and Sample Distribution in Telephone Surveys

Singer, et al (1999a) found incentives generally improve response rates in telephone surveys, and these effects do not differ by interviewing mode. Other relevant effects on survey participation in telephone surveys include (1) an increase in the response difference between incentive and no incentive conditions for high burden surveys and (2) a greater impact on the response rate for surveys with relatively low participation without the incentives. The first finding suggests that longer and/or more demanding surveys benefit more from incentives than shorter and/or less demanding surveys. The second finding indicates that incentive payments are quite effective when other motives to participate are absent, as Singer (2002) points out.

The effectiveness of incentives for improving overall response rates in both routine and more burdensome surveys suggests that paying incentives to initial refusers would generally produce a more representative sample. All else being equal, the lowered nonresponse rate is likely to reduce the potential (or degree) of nonresponse bias. Of course, this conclusion only holds if the increased response decreases differences between respondents and nonrespondents. In theory, increasing participation is likely to reduce nonresponse bias, but for some samples greater response may have no effect on sample composition or actually produce a less representative sample. For example, if respondents who initially refused have similar characteristics and give similar answers to those who did not refuse, adding these sample members to the final sample will not have a significant impact on sample representativeness.

An important point about offering an incentive payment to initial refusers is that this practice generally increases the participation of sample members who would otherwise be difficult to include in the survey (Singer, et al, 1999a). For example, smokers may be more reluctant than non-smokers to participate in a survey about smoking sponsored by their state health department or a federal health agency. Smokers may be unhappy with increasing limitations on public smoking, they may be wary of being admonished about their smoking, or they may have other reservations about participating in a study of smoking behavior and attitudes. In this scenario, refusal conversion payments would seem likely to produce a more representative

sample than would otherwise be the case by including more smokers in the final sample. This outcome would represent a major intended consequence of refusal incentives on sample composition – that is, to improve the representativeness of the sample by including those who are more difficult to involve in the survey.

Refusal conversion payments might also have unintended, negative consequences on the composition of the final sample. For example, a survey on smoking behavior might offer incentives to initial refusers who have relatively extreme views on tobacco use, such as being against any and all limits on smoking in any public places. The refusal incentives might therefore produce data that somewhat over-represents members of the population with extreme positions, at the expense of more representative views.

Refusal Conversion Incentives and Data Quality in Telephone Surveys

Survey researchers have also recently been concerned about another unintended consequence of paying incentives – the impact of incentives on the quality of data paid respondents provide. Groves and his colleagues recently found some evidence that incentives are effective in persuading sample members who have little interest in the survey topic to participate (Groves, Singer, and Corning, 2000), although more recent experiments produced mixed results on the consistency of this relationship (Groves, Presser, and Dipko, 2004). The primary concern of survey researchers is that respondents with little interest in the survey topic might also be less motivated to provide complete and accurate data than other sample members (Shettle and Mooney, 1999; Singer, 2002). As a result, respondents who are motivated to participate in a survey primarily by an incentive might not meet the same quality standards as respondents motivated by other factors, such as interest in the topic or sponsor, or a sense of fulfilling a civic duty (Groves, et al, 2000).

To date, little evidence has indicated that incentives have a negative impact on data quality. In fact, research conducted to date have shown either (1) no differences in data quality between those who were paid an incentive and those who were not or (2) that paid respondents actually provided somewhat higher quality data than those who were not paid (Singer, et al, 1999a). Yet none of the studies comparing data quality

directly compared respondents who were paid incentives after refusing an earlier request to participate with unpaid non-refusers. Apparently, little research exists on the impact of refusal conversion incentives on the quality of survey responses. Sample members who initially refuse participation in a survey very commonly express a lack of interest in the survey. A comparison of data quality between paid and unpaid respondents based on refusal status is therefore likely to be similar to a comparison between more and less interested respondents. In surveys that offer only refusal conversion incentives, the relationship between incentive status and data quality might prove more relevant.

Research Questions

This study focuses on three research questions to assess the impact of a survey protocol offering incentives to initial refusers on sample composition and data quality:

1. Do those who initially refused to participate differ significantly from those who never refused in terms of key demographic characteristics?
2. Do initial refusers and non-refusers differ significantly on the main substantive survey items, including tobacco and health indicators?
3. Do initial refusers and non-refusers differ significantly on the response effort they provided, including inability to give a valid response and refusal to respond?

If almost no significant differences between initial refusers and non-refusers are observed on these indicators, then we will be led to conclude that the refusal incentives had no substantial impact on the data collection results, beyond the intended effect of increasing the overall response rate. If significant differences between initial refusers and non-refusers are observed, then we will have some evidence that refusal incentives did have a substantial impact on the survey estimates. Finding significant differences between initial refusers and non-refusers with respect to sample composition will prompt further analysis to determine whether refusal incentives produce intended or unintended consequences, or some combination of the two. To address these issues, this study compares data between initial refusers and non-refusers from the 2003 NY ATS.

Methods

Survey Data

The New York Adult Tobacco Survey (ATS) is a statewide household survey of New York residents aged 18 and older that began with the third quarter of 2003. This 25 to 30 minute survey measures the prevalence of use of smoking, smoking behaviors, and attitudes towards smoking and smoking cessation efforts. These quarterly surveys are designed to provide timely surveillance and evaluation data to inform the New York state tobacco control program.

The NY ATS sample follows a stratified dual frame design, with sampled telephone numbers drawn equally from (1) a random-digit-dial (RDD) frame and (2) a residential directory-listed frame. This design provides a representative sample while increasing the “hit rate” of current residential units to improve data collection efficiency. Both sample frames are stratified across eight county groups – the Buffalo Area, Hudson Valley Area, Northern Capital Area, Northern Central Area, New York City-Long Island Area, Rochester Area, Southern Capital Area, and Southern Central Area – producing a total of 16 strata. The sample design permits the calculation of precise statewide and regional estimates.

Within each household associated with the sampled numbers, one adult age 18 or older is selected. The CATI selection protocol also oversamples current smokers by selecting 100% of current smokers identified during the initial household listing of adults and smokers. Prior to weighting, this oversampling of smokers within households increases their proportion in the final sample about 30% over the expected proportion in the state.

A key feature of the NY ATS survey protocol for the present study is that interviewers do not initially offer selected sample members an incentive payment to participate. When selected sample members, or other members of the household, refuse the initial attempt to recruit one adult into the study, the protocol specifies an offer of a \$20 cash incentive to refusers on the next contact. For each quarter of the NY ATS, initial refusers who complete the survey after being offered the incentive payment comprise about 25% of the final sample.

The present study uses NY ATS data from quarter 4 of 2003. A total of 2,063 interviews were completed in this quarter. Of these, 1,500 (72.7%) had not ever refused to participate and 563 (27.3%) had initially refused to participate. The AAPOR4 response rate for the quarter was 23.8%. The average interview length was 26.7 minutes. On average, the survey is about 6 minutes longer for current smokers (31 minutes) compared to non-smokers (25 minutes). Overall, the data collection results from this quarter look fairly consistent with other quarters.

Measures and Analysis

Quarter 4 of the 2003 NY ATS included measures needed to analyze the impact of the refusal incentive protocol on sample composition and data quality. Standard demographic items such as respondent's age, gender, ethnicity, race, education, income, employment status, and marital status were measured. The main substantive lines of questioning included tobacco use, beliefs about the health risks of smoking, views on current state smoking policies, exposure to anti-smoking media campaigns, and current health status and conditions. Among all of these survey items, the analysis used current smoking status, two beliefs about smoking items, two exposure to anti-smoking campaigns items, support for a statewide ban on public smoking, and indicators of current health. The Appendix provides each of the specific survey items and the coding of each variable in the analysis.

The analysis required to address the three main research questions is straightforward. To provide evidence to assess the first question, the appropriate analysis comprised independent samples t-tests between initial refusers and non-refusers for each dichotomized demographic variables. For the second research question, the analysis involved t-tests between initial refusers and non-refusers for the dichotomized substantive variables including: current smoking status, beliefs about the dangers of smoking, exposure to anti-smoking messages on television and in print, support for the statewide ban on public smoking, current health status, and health-related limitations. The analysis for the third research question followed this same approach at the other analyses, using t-tests between initial refusers and non-refusers on the mean number of "Don't Know" responses given, the mean number of refused

items, and the mean interview length. All of these t-tests used unweighted data and assumed unequal variances. Statistical significance was set at the traditional cut-off, where p-values less than .05 are considered significant.

Results

Demographic Differences between the Refusal and Non-refusal Groups

To address the first research question, analysis compared refusers and non-refusers on several demographic characteristics. Table 1 provides the results of this analysis, based on independent sample t-tests with unweighted data. Initial refusers and non-refusers differed on only three of the nine demographic characteristics examined – age distribution, education level, and employment status. Those who initially refused to participate in the NY ATS (mean age 49.4) were, on average, somewhat older than participants who did not refuse (mean age 47.4). Among refusers, 41.9% were under age 45 while 46.3% of non-refusers were under 45. On the higher end of the range, 38.0% of refusers were age 55 and over, compared to only 32.6% of non-refusers. This finding is consistent with Shettle and Mooney's (1999) comparison of incentive and non-incentive groups among retired Current Population Survey (CPS) sample, which found that incentives had a greater impact on the older respondents than the younger ones. Refusers were also somewhat less educated than non-refusers. While 35.2% of refusers had completed a 4 or 5-year college degree, 41.6% of non-refusers held a college degree. Initial refusers were also less likely to be employed than non-refusers, and more likely to be homemakers, students, retired, or otherwise not currently employed for pay. Only 52.3% of refusers were currently employed for pay, but 63.8% of refusers were employed. No significant differences were observed on the other six demographic indicators, although the non-significant difference between refusers and non-refusers on income level was consistent with the significant differences observed on college completion and employment status.

Table 1: Comparison of Demographic Characteristics by Refusal Conversion Status in the New York Adult Tobacco Survey

Demographic Characteristics (unweighted data)	No Refusal (n = 1,500)	Initial Refusal (n = 563)
Age *		
18 – 24	7.2 %	8.2 %
25 – 34	17.4	14.1
35 – 44	21.7	19.6
45 – 54	21.0	20.1
55 – 64	15.2	16.3
65 +	17.4	21.7
Female	61.4 %	61.5 %
Hispanic/Latino background	5.4 %	6.4 %
Non-White race	19.3 %	20.2 %
College degree *	41.6 %	35.2%
Currently employed for pay *	63.8 %	52.3 %
Currently married	50.1 %	49.2 %
Children in household	36.5 %	38.1 %
Income \$30,000 or more	70.3 %	66.7 %

* Difference between No Refusal and Initial Refusal cases is statistically significant at $p < .05$ based on independent samples t-tests.

Smoking and Health Differences between the Refusal and Non-refusal Groups

To address the second research question, analysis compared refusers and non-refusers on several substantive indicators dealing with tobacco use and health status. Table 2 provides the results of this analysis, based on independent sample t-tests with unweighted data. Overall, initial refusers and non-refusers differed on only two of the eight substantive indicators examined. Yet one of the indicators where a significant difference was observed was a central variable in the NY ATS, current smoking status. Unweighted, 24.8% of those who did not initially refuse reported current smoking while

Table 2: Comparison of Smoking and Health Indicators by Refusal Conversion Status in the New York Adult Tobacco Survey

Smoking and Health Indicators (unweighted data)	No Refusal (n = 1,500)	Initial Refusal (n = 563)
Current Smoker *	24.8 %	28.9 %
Believe little benefit to quitting if smoke pack a day for 20 years	23.4 %	23.3 %
Believe better to smoke light cigarettes if have to smoke	48.4 %	47.3 %
In favor of state law banning smoking in all public places	71.8 %	72.4 %
Recently seen anti-smoking advertisements on television at least once per week	63.0 %	63.2 %
Recently seen anti-smoking advertisements in newspaper or magazines at least once per week	41.1 %	39.0 %
Overall health very good/excellent	58.7 %	55.5 %
Health limits activities *	18.3 %	23.1 %

* Difference between No Refusal and Initial Refusal cases is statistically significant at $p < .05$ based on independent samples t-tests.

28.9% of initial refusers were smokers. Clearly, the refusal conversion incentive pulled more smokers into the sample than would have been the case if they had not been offered \$20. Interestingly, despite this difference in the proportion of current smokers, the refusal and non-refusal groups did not differ in their responses to the other tobacco-related items, including beliefs on the health risk of smoking and support for the statewide ban on smoking in public places. The only other substantive indicator on which the refusers and non-refusers differed was having a health condition that currently limits their daily activities. Only 18.3% of non-refusers reported such a limitation, while 23.1% of initial refusers reported this situation. This finding seems logical, in that those sample members with significant health problems are likely to be

more difficult to include in telephone surveys. These sample members may be more likely than others to view the incentive payment provides as a necessary compensation for the burden associated with completing a 25 to 30 minute survey.

Further Analysis of Differences between the Refusal and Non-refusal Groups

The observed differences between initial refusers and non-refusers on three demographic characteristics and current smoking status warrant further analysis. Although there were very few differences between the refusal and non-refusal groups overall, the observed differences are quite relevant to the NY ATS. Clearly, smoking status is a critical indicator in the ATS. The three demographic items – age, education, and employment status – are not quite as central to the survey goals, but each could still have an important impact on the representativeness of the sample. For these reasons, it's useful to know whether the refusal incentives tend to produce a more or less representative final sample. Evidence on the representativeness of the final sample will indicate whether using refusal conversion incentives appears to decrease or increase non-response bias.

We therefore need to add another research question: do the demographic and substantive data for the refusal group versus the non-refusal group appear to be more representative of the population, based on estimates from similar statewide telephone surveys? To answer this question, a comparison of the NY ATS demographic and smoking status data for refusers and non-refusers was made with parallel estimates from the 2003 New York Behavioral Risk Factors Surveillance Survey (BRFSS). The New York BRFSS is an ongoing annual survey sponsored by the Centers for Disease Control that collects data on behaviors related to health risk factors, including smoking. It is important to note that the BRFSS is not necessarily a “gold standard” against which we can measure the representativeness of health indicators or demographic composition in the NY ATS. Determining the precise degree of non-response bias in the ATS data is not the intent of this comparison. Rather, the point is that the BRFSS is another statewide, RDD-based telephone survey on health behaviors, like the NY ATS. We can therefore use the BRFSS data as a realistic benchmark to assess whether the data from the refusal group versus the non-refusal group look more similar to the data

Table 3: Comparison of Demographic Characteristics and Smoking Indicators between the New York ATS (Q4 2003) and the New York BRFSS (2003)

Demographic and Smoking Indicators (weighted data)	NY ATS Q4 '03 No Refusal (n = 1,500)	NY ATS Q4 '03 Initial Refusal (n = 563)	NY BRFSS 2003 (n = 5,472)
Age *			
18 – 24	12.2 %	13.4 %	12.5 %
25 – 44	42.2	34.3	38.7
45 – 64	30.3	29.2	31.2
65 +	15.1	23.1	17.6
Gender			
Male	47.4 %	46.7 %	47.4 %
Female	52.6	53.3	52.6
Race/ethnicity			
Non-Hispanic White	64.6 %	59.9 %	62.6 %
Non-Hispanic Black	14.2	12.6	12.2
Hispanic/Latino	12.7	16.7	16.6
Other race/ethnicity	8.5	10.8	8.7
Education *			
Less than High School	5.8 %	8.6 %	13.1 %
High School or equivalency	28.9	29.6	29.7
Some college	24.1	25.0	23.6
College degree or higher	41.2	36.8	33.5
Income			
Less than \$10,000	4.2 %	3.1 %	6.7 %
\$10,000 to \$20,000	7.7	11.1	14.6
\$20,000 to \$50,000	38.6	38.4	38.0
\$50,000 or more	49.5	47.4	40.7
Smoking Status *			
Current Smoker	20.3 %	27.3 %	21.6 %
Non-smoker	79.7	72.7	78.4

* Difference between NY ATS No Refusal and Initial Refusal cases is statistically significant at $p < .05$ based on independent samples t-tests.

we would expect to observe using the same survey methods for this population., namely the state of New York.

Table 3 shows comparisons among data the non-refusal group the refusal group from quarter 4 of the 2003 NY ATS and the 2003 NY BRFSS estimates. Parallel measures from the NY ATS and the NY BRFSS were available for five demographic characteristics and current smoking status. As necessary, the data for the demographic items were recoded to match the readily available categorical data from the 2003 NY BRFSS. In addition, this comparison uses weighted NY ATS data, since the available BRFSS data are likewise weighted. Another important methodological note is that the BRFSS does not offer any incentives to respondents, regardless of refusal status.

Starting with the three measures that were significantly different between ATS refusers and non-refusers, the age group estimates for the non-refusers appear to be more similar to the BRFSS estimates. Both the refusal and non-refusal ATS groups differ considerably from the BRFSS estimate in the age 25 to 44 category (in opposite directions). Yet estimates for the other three other age categories look more similar between the non-refusers and BRFSS than between the refusers and the BRFSS.

With respect to education, both ATS groups have similarly lower estimates of New York residents who have not completed high school compared to the 2003 BRFSS. Interestingly, all three groups have similar proportions of respondents with high school diploma (or equivalency) and some college. Considerable differences are observed between the three groups in the college degree or higher group. In this category, both ATS groups provide higher estimates than the BRFSS. Yet the refusal group is considerably more similar to the BRFSS sample than the non-refusal group. Overall, the non-refusal group provides the most educated sample, a sample which appears to over-represent respondents with a college degree or higher education.

Smoking status is clearly the most salient difference observed between initial refusers and non-refusers on the ATS. Comparing smoking status for these two groups with the BRFSS indicates that the non-refusal group is considerably more similar than then refusal group. The data suggest the ATS refusers significantly over-represent the proportion of smokers in the New York state population.

Although none of the other variables compared across the ATS and BRFSS surveys had significant differences between ATS refusers and non-refusers, a few additional comparisons are noteworthy. All three groups had similar proportions of male and female respondents, although the NY ATS refusal group had slightly fewer women. With respect to race and ethnicity, all three groups produced similar estimate. One difference was that the non-refusal ATS group had slightly more non-Hispanic Black respondent, but slightly fewer respondents with a Hispanic/Latino background. Also, the ATS refusal group had slightly more respondents grouped in the other race/ethnicity category than the ATS non-refusal group or the BRFSS sample. A striking difference between the two ATS groups and the BRFSS sample was that both groups included fewer respondents with incomes below \$20,000. At the other end of the scale, both ATS groups also had higher proportions of respondents with incomes greater than \$50,000. Combined with the results for education level, the findings for income level indicate that the ATS appears to under-represent those with lower socioeconomic status (SES) and over-represent those with higher SES compared to the BRFSS. Of course, even more representative New York estimates from a source such as the Census could resolve absolute differences in representation more definitively.

Differences in Data Quality between the Refusal and Non-refusal Groups

The third research question involved a different issue from the first two, that is, potential differences in data quality between the refusal and non-refusal groups. To address this research question, analysis compared refusers and non-refusers on three basic indicators of the response effort exerted: number of “don’t know” answers given, number of items refused, and overall duration of the interviews. Table 4 provides the results of this analysis, based on independent sample t-tests with unweighted data. Initial refusers and non-refusers differed on all three of these indicators, but not in a consistent way. First, the non-refusal group (mean = 3.56 items) actually rendered more don’t know answers than the refusal group (mean = 3.14 items). This indicates that the non-refusal group was more likely to fail to choose a valid response option for the survey items they answered. Yet refusers (mean = 0.66 items) declined to answer

Table 4: Comparison of Response Effort by Refusal Conversion Status in the New York Adult Tobacco Survey

Response Effort Indicators (unweighted data)	No Refusal (n = 1,500)	Initial Refusal (n = 563)
Mean number of Don't Know responses *	3.56	3.14
Mean number of Refused items *	0.48	0.66
Mean Interview Length in minutes *	26.4	27.5

* Difference between No Refusal and Initial Refusal cases is statistically significant at $p < .05$ based on independent samples t-tests.

more survey items than non-refusers (mean = 0.48 items). So, while non-refusers were more likely to provide “don’t know” answers to survey items than refusers, refusers were more likely to refuse to answer questions compared to non-refusers.

Another potential indicator of overall response effort is the total length of the interview. Although most respondents tend to follow the pace set by the interviewer and attempt to answer questions as well as they can, some respondents may attempt to lower the survey burden by rushing through the interview as quickly as they can, without concern for the accuracy or completeness of their responses. Mean comparisons show that refusers (mean = 26.4 minutes) had a significantly greater interview length than non-refusers (mean = 27.5 minutes). While this difference is statistically significant, this finding is likely explainable by the simple fact that the refusal group included a higher proportion of current smokers. The NY ATS interview is about 5 to 6 minutes longer for current smokers than non-smokers. With slightly more than 25% of the refusal group being current smokers, the additional 1 minute of interviewing time compared to the non-refusal group seems appropriate. Hence, the difference in interview length does not likely tell us anything meaningful about the overall survey effort exerted by refusers and non-refusers.

Discussion

The goal of this study was to assess the impact of using refusal conversion incentives on sample composition, distribution of responses, and data quality in a telephone survey. Using data from the NY ATS, we compared those who initially refused to participate with those who did not initially refuse on key demographic and substantive items and indicators of data quality. The results indicated that the refusal conversion incentives increased the proportion of respondents who were over age 55, did not have a college degree, and were not currently employed for pay. Furthermore, those who initially refused were more likely to be current smokers and to have a health condition that limited their daily activities. The results for indicators of data quality were mixed: the refusal group was less likely to provide “don’t know” answers than the non-refusal group, but refusers were more likely to decline to answer items than the non-refusers. Overall, these findings provide mixed evidence on the intended and unintended consequences of using differential incentive schemes such as refusal conversion payments.

First, even though there were only a few demographic differences between the refusal and non-refusal groups in the NY ATS, the observed differences indicated response some bias toward more educated and employed sample members without the incentive payments. Using the \$20 refusal conversion incentives pulled in more sample members who did not have a college degree and were not currently employed for pay. Comparisons among the two NY ATS groups and 2003 NY BRFSS estimates reinforced this tendency in the ATS data. This evidence supports a preliminary conclusion that offering incentives to initial refusers entices less educated respondents to participate and thereby improves the overall representativeness of the sample on this dimension. This outcome reflects one of the major intended consequences of using refusal conversions incentives, to include sample members who are generally less likely to participate in the survey.

Second, the refusal conversion incentives were more salient to current smokers, producing an over-representation of smokers in the final NY ATS sample. This outcome is both intended and positive, on one hand, but also somewhat unintended and negative on the other. On the positive side, the inclusion of more smokers in the sample through

paying initial refusers allows for more rigorous comparisons between smokers and non-smokers in the NY ATS sample using unweighted data. Furthermore, to the extent that smokers may be more reluctant than non-smokers to participate in this kind of survey, this outcome conforms to the intended consequence of pulling in more sample members who are less likely to participate. Yet the obvious unintended consequence is even further over-representation of smokers in the final sample. Comparing the weighted NY ATS data to the NY BRFSS data reinforces this point, in that the proportions of current smokers for the ATS non-refusal group and the BRFSS sample are 5% to 6 % lower than the ATS refusal group. Furthermore, when the NY ATS data are combined and weighted, the added value of including the additional smokers becomes minimal, since these cases need to be weighted down.

Third, consistent with most existing studies of the impact of incentives on data quality, the NY ATS findings did not demonstrate a clear pattern of lower data quality in the responses provided by the refusal group. Initial refusers were more likely to provide valid answers for the questions that they did agree to answer, but on average they declined to answer slightly more questions than the non-refusers. Overall, these data did not exhibit the potential unintended consequence of lower data quality among initial refusers who were provided incentive payments to participate. More extensive analysis of data quality issues, such as invariance of responses to items on common scales (such as agree-disagree items) and less complete responses to open-ended items, could provide a more definitive picture of the response effort applied by refusers and non-refusers.

This research can be extended in at least a few ways to provide more compelling evidence on how refusal conversion incentives affect sample composition, distribution of responses, and data quality. To start, replicating the analysis with additional quarters of NY ATS data or pooling data across survey quarters (for example, by year) will confirm whether the findings for Q4 of 2003 are representative or not. Some variation in the composition and responses of initial refusers is certainly possible across quarters, even when the survey methods remain virtually identical. It would also be useful to include a greater range of substantive items in the analysis. This expansion of the analysis would provide a more comprehensive assessment of the overall impact of paying initial

refusers on the distribution of survey responses. Expanding the measures of data quality would provide a more thorough evaluation of the impact of refusal conversion incentives on the response effort exerted by initial refusers. Although this information is somewhat limited, additional indicators could include consistency of responses to items using the “agree-disagree” answer format and completeness of open-ended responses.

Another useful elaboration of this research would be to evaluate the impact of refusal conversion incentives on subsequent participation in a similar survey. Research on differential incentives by Groves and Singer and their colleagues found that a clear majority of respondents felt it was unfair for survey organizations to pay uncooperative sample members to participate in the same survey (Groves, et al, 1999; Singer, et al, 1999b). Interestingly, these feelings of inequity did not appear to translate into greater reluctance to participate among respondents in a subsequent survey wave (Singer, et al, 1999a). Furthermore, the effects of receiving an incentive at wave 1 and refusing the earlier survey on participation at wave 2 were not statistically significant. In the case of NY ATS respondents, those who were currently smoking or had recently quit smoking when they originally participated in the ATS are re-contacted in about one year for a follow-up survey. For the follow-up ATS, the incentive procedures are the same as the original cross-sectional ATS – sample members are not offered an incentive until they refuse to participate. The follow-up survey will therefore allow us to add another dimension to research on differential incentives by looking at whether former refusers behave the same way when presented with a follow-up survey opportunity. The question is whether sample members who initially refused to participate in a survey and were then offered an incentive payment to complete the survey thus learn to use the refusal strategy again to secure payment for the follow-up (Singer, et al, 1999b). This further research will provide data on a further unintended consequence of differential incentive plans – essentially training sample members to “hold out” until a monetary incentive is offered. As the preliminary investigation of the impact of refusal conversion incentives has provided evidence of both intended and unintended consequences, more thorough examination of the outcomes associated with these kinds of incentive plans seems warranted. fulfill

References

- Church, Allan H. (1993). "Estimating the Effect of Incentives on Mail Survey Response Rates: A Meta Analysis." *Public Opinion Quarterly*, 57:62–79.
- Groves, Robert M., Stanley Presser, and Sarah Dipko. 2004. "The Role of Topic Interest in Survey Participation Decisions." *Public Opinion Quarterly*, 68:2–31.
- Groves, Robert M., Eleanor Singer, Amy D. Corning, and Ashley Bowers. 1999. "A Laboratory Approach to Measuring the Effects on Survey Participation of Interview Length, Incentives, Differential Incentives, and Refusal Conversion." *Journal of Official Statistics*, 15:251–268.
- Groves, Robert M., Eleanor Singer, and Amy D. Corning. 2000. "Leverage-Saliency Theory of Survey Participation: Description and an Illustration." *Public Opinion Quarterly*, 64:299–308.
- Shettle, Carolyn and Geraldine Mooney. 1999. "Monetary Incentives in U.S. Government Surveys." *Journal of Official Statistics*, 15:231–250.
- Singer, Eleanor, John Van Hoewyk, Nancy Gebler, Trivellore Raghunathan, and Katherine McGonagle. 1999a. "The Effect of Incentives on Response Rates in Interviewer-Mediated Surveys." *Journal of Official Statistics*, 15:217–230.
- Singer, Eleanor, Robert M. Groves, and Amy D. Corning. 1999b. "Differential Incentives: Beliefs about Practices, Perceptions of Equity, and Effects on Survey Participation." *Public Opinion Quarterly*, 63:251-260.
- Singer, Eleanor. 2002. "The Use of Incentives to Reduce Nonresponse in Household Surveys." In *Survey Nonresponse*, ed. Robert M. Groves, Don A. Dillman, John L. Eltinge, and Roderick J.A. Little, pp.163-177. New York: Wiley.

Appendix: Demographic Characteristics and Tobacco and Health Indicators from the 2003 New York Adult Tobacco Survey

Demographic Characteristics	New York ATS Survey Item	Item Coding
Age	What is your age? _____ (AGE IN YEARS)	1 = 18-24 2 = 25-34 3 = 35-44 4 = 45-54 5 = 55-64, 6 = 65+
Hispanic/Latino background	Are you Hispanic or Latino? 1. Yes 2. No	0 = no 1 = yes
Race	Which one or more of the following would you say is your race: 1. White 2. Black or African American 3. Asian 4. Native Hawaiian or Pacific Islander 5. American Indian, Alaska Native 6. Other (specify)	0 = white 1 = other than white
Gender	For survey purposes, I need to confirm if you are male or female? 1. Male 2. Female	0 = male 1 = female
Education level	What is the highest level of school you completed or the highest degree you received? 1. Never attended school/only attended kindergarten 2. Grades 1 through 8 (Elementary) 3. Grades 9 through 12 (Some high school) 4. Grade 12 (High school graduate) 5. G.E.D 6. Some technical or vocational school 7. Some college, no degree 8. AA; technical or vocational school 9. AA; academic 10. BA, BS (College graduate) 11. At least some graduate or professional school 12. Graduate or professional degree	0 = less than college degree 1 = college degree or higher
Employment status	Are you currently... 1. Employed for wages 2. Self-employed 3. Out of work for more than 1 year 4. Out of work for less than 1 year 5. Homemaker 6. Student 7. Retired 8. Unable to work	0 = not currently employed 1 = employed/self-employed

Income level	Was your annual household income from all sources during 2002 more or less than \$30,000? 1. \$30,000 or more 2. Less than \$30,000	0 = income greater than \$30K 1 = income less than \$30K
Marital status	Are you: 1. Married 2. Divorced 3. Widowed 4. Separated 5. Never married, or 6. Living with a partner	0 = not currently married 1 = currently married
Children	How many children live in your household who are... ... younger than 5 years old? _____ ... 5 through 11 years old? _____ ... 12 to 17 years old? _____	0 = no children in household 1 = at least one child in household
Tobacco and Health Indicators	New York ATS Survey Item	Variable Coding
Smoking status	a. Have you smoked at least 100 cigarettes in your entire life? 1. Yes 2. No b. Do you now smoke cigarettes everyday, some days, or not at all? 1. everyday 2. some days 3. not at all	0 Non-smoker: a = 2 or (a = 1 and b = 3) 1 Smoker: a = 1 and (b = 2 or b = 3)
Belief in value of quitting after prolonged smoking	If a person has smoked a pack of cigarettes a day for more than 20 years, there is little health benefit to quitting smoking. 1. Strongly agree 2. Agree 3. Disagree 4. Strongly disagree	0 = strongly disagree or disagree 1 = agree or strongly agree
Belief in value of smoking light cigarettes	If you have to smoke, you are probably better off smoking a light cigarette. 1. Strongly agree 2. Agree 3. Disagree 4. Strongly disagree	0 = strongly disagree or disagree 1 = agree or strongly agree

<p>Exposure to anti-smoking ads on television</p>	<p>In the past 30 days, how often have you noticed anti-smoking advertising or information on television?</p> <ol style="list-style-type: none"> 1. 2-3 times per day 2. Every day 3. Once per week 4. Less than once per week 5. Never 	<p>0 = never or less than once per week, 1 = once per week, every day, or 2-3 times per day</p>
<p>Exposure to anti-smoking ads in newspapers or magazines</p>	<p>In the past 30 days, how often have you noticed anti-smoking advertising or information on television?</p> <ol style="list-style-type: none"> 1. 2-3 times per day 2. Every day 3. Once per week 4. Less than once per week 5. Never 	<p>0 = never or less than once per week, 1 = once per week, every day, or 2-3 times per day</p>
<p>Support for New York state smoking ban in public places</p>	<p>Are you personally in favor, opposed to, or indifferent to the recently enacted New York State law prohibiting smoking in all public and work places, including bars and restaurants?</p> <ol style="list-style-type: none"> 1. favor 2. oppose 3. indifferent 4. not familiar with law 	<p>0 = oppose or indifferent 1 = favor</p>
<p>Overall health status</p>	<p>In general, would you say your health is:</p> <ol style="list-style-type: none"> 1. Excellent 2. Very good 3. Good 4. Fair 5. Poor 	<p>0 = poor, fair, or good, 1 = very good or excellent</p>
<p>Health limitation</p>	<p>Are you limited in any way in any activities because of physical, mental, or emotional problems?</p> <ol style="list-style-type: none"> 1. Yes 2. No 	<p>0 = no 1 = yes</p>