Population Surveys from RDD Telephone and Internet Panel Samples:
A Weighted Comparison of Two National Taxpayer Surveys

Paper presented at the 2005 Federal Committee on Statistical Methodology Research Conference

John M. Boyle, Ph.D.
Schulman, Ronca & Bucuvalas, Inc.
8403 Colesville Road, Silver Spring, MD 20910

George Freeland & Lynda Mulvany
Internal Revenue Service
1111 Constitution Avenue, Washington, D.C. 20224

Abstract

A nationwide telephone survey of 2,000 randomly-selected households was conducted in January-February 2004 to determine taxpayer use and perceptions of the paper and electronic versions of Publication 17 and the IRS website from which the electronic version can be obtained. The survey was conducted for the Tax Forms and Publications (TFP) Division of the Internal Revenue Service (IRS) by Schulman, Ronca, and Bucuvalas, Inc. (SRBI). This telephone survey provides estimates of taxpayer characteristics, experiences, attitudes and preferences based upon a national probability sample.

As part of the qualitative research phase of this project, a web-based survey was also conducted among a national Internet panel. A total of 9,000 invitations were e-mailed to a national sample of adult panelists. Nearly 1,000 qualified responses were received from taxpayers to the Internet survey. The response rate was reported as standard for a web-based survey of this length by the organization who managed the panel.

Since many of the questions were identical for the telephone and Internet samples of taxpayers, a comparison of the findings between the two surveys should illuminate the implications of these two modes of data collection for surveys of taxpayers. Since approximately ninety percent of adults in the United States meet the study definition of taxpayers, these findings should also be useful for comparisons to other national population surveys.

For this paper, both samples have been restricted to the person in the household most familiar with 2002 federal income tax, and the Internet sample was weighted to the parameters of the telephone sample on age, gender and education to correct for known population biases in Internet surveys. Nonetheless, many significant differences persist in other characteristics, attitudes and behaviors of the weighted Internet sample and the national telephone sample. Some of these differences may enhance the value of Internet panels in qualitative research, but reinforce existing concerns about its utility for population estimates.

Research Issue

Approximately twenty years ago, telephone interviewing among samples selected by random digit dialing replaced face-to-face interviewing among samples selected by area probability sampling as the dominant mode of surveys in both the private and public sectors in the United States. The proportion of households in the U.S. with telephone service was 93% in 1980 (compared to 87% in 1970). The limited coverage bias associated with a telephone household frame was felt to be offset in most instances by the speed, lower costs and better quality control (interview monitoring) associated with the mode of interviewing, and the reduced design effects on sampling estimates. In addition, the increasing difficulty of gaining access to systematically selected dwelling units for face to face interviews was raising questions about the quality of the effective sample for most in-person surveys.

Today, a similar although not equivalent situation confronts telephone surveys. The household and population coverage rates for telephone surveys are being eroded by cell phone only households. Answering machines, voice mail, and call managers are increasing the difficulty in reaching individuals in households with landline telephones. Refusal rates are rising and the "do not call" lists may be seen by the public as legitimizing refusals to surveys, as well as telemarketing.

At the same time, web-based surveys of Internet panels offer a lower cost, often quicker data collection option to telephone. This alternative mode of data collection has been embraced by large segments of the commercial world for reasons of both cost and speed. It has also been embraced by many market research firms because it allows them to eliminate much of the data collection infrastructure associated with telephone interviewing, particularly the recruitment and training costs of telephone interviewers in response to shifts in workload.

Although the pressures for a shift in survey mode from telephone to Internet in the first part of the Twenty First Century are similar to the pressures for a shift from face-to-face to telephone that occurred two decades ago, there are a couple of big differences in terms of general population surveys. First, the household penetration rate for Internet access is only about sixty-five percent today, compared to over ninety-percent when telephone households were accepted as an acceptable sampling frame for the U.S. population. Second, there is no national sampling frame for Internet households that could be adapted to random or systematic sampling such as the ten digit telephone number (with or without banks with listed numbers). Third, the laws concerning unsolicited e-mails by Internet, and the technology available for avoiding such contact, are far more extensive for Internet surveys compared to telephone surveys.

Consequently, national population surveys by Internet in the United States at the present time are based on pre-recruited panels of households and/or individuals who have Internet access. In some instances, the initial recruitment of the household or individual may be probability based. In some instances, the selected household may be

offered Internet access in return for participation in the survey(s). However, the most common sampling frame for these surveys is Internet panels that are recruited at Internet portals and websites.

There are significant costs associated with the recruitment and maintenance of a large Internet panel. Hence, most Internet panels invite panelists to participate in a large number of surveys on a continuous basis. At the same time, in order to maintain the panel, the companies who offer these panels do not want to annoy panelists by pressuring them to participate in surveys that don't interest them. Hence, the most common approach for an Internet panel survey is to send an initial invitation to a sufficiently large sample of panelists to obtain the targeted number of completed interviews without the use of follow-up invitations or reminders.

Those who argue that valid population estimates can be obtained for web-based surveys of Internet panels usually raise demographic weighting procedures to correct for sampling bias in these surveys. Certainly the well known age and income biases associated with Internet access at home can be adjusted by sample weighting. But does this adjustment of a limited set of sample characteristics to population characteristics actually correct the biases in an Internet sample that are associated with self-selection for the panel and self-selection for the specific survey, as well as coverage biases?

Two surveys conducted by SRBI for the Internal Revenue Service provide an opportunity to test the hypothesis that bias in sample estimates in a national survey conducted by Internet panel can be corrected by demographic weighting. The first survey was a web-based survey conducted among a national Internet panel. The second survey was conducted as a telephone survey among a national RDD sample of households.

Background to the Surveys

The two surveys described in this report were conducted for Tax Forms and Publications of the Internal Revenue Service as part of an evaluation of customer satisfaction with IRS Publication 17, Your Federal Income Tax. Publication 17 is a comprehensive tax guide for individuals that contains the Service's most comprehensive discussion of the provisions of Federal income tax law that affect the vast majority of individuals in the country. It is the largest publication and the most widely used of the technical information publications issued by the Service.

Publication 17 is available in a hard copy format and is also available in Adobe PDF format on the Service's web site, www.irs.gov. Material in the publication is searchable via the Adobe search engine and through the web site's search engine. The division is considering redesigning the electronic version of this publication, in part, to increase taxpayer use of the electronic version rather than the paper version. Before attempting to redesign the electronic version of the publication, the IRS wanted to have a better understanding of how customers use this product and listen to their suggestions on design and topic areas needing change and improvement. Focus groups were conducted among users of hard copies of Publication 17 and electronic forms of Publication 17. However, there had been no previous large-scale quantitative assessment of taxpayer knowledge and experience with Publication 17 in either format. To further its knowledge of customer needs, customer usage traits, and customer satisfaction, TFP commissioned a survey of a national sample of taxpayers to explore their awareness, use, perceptions and preferences for both the paper and the electronic version of Publication 17. This population survey provides a context for the earlier qualitative research on Publication 17, including a national qualitative survey conducted by Internet.

Survey Design

The primary objective of the survey was to obtain insights into taxpayer **perceptions, experiences, and satisfaction** with the paper and electronic version of Publication 17, as well as the IRS website on which the latter is hosted. Specifically, the survey explored:

- Experience and satisfaction with locating and searching the publication;
- Ease of use of the publication;
- Whether the use of the publication accomplished the desired result (e.g., was the sought answer found?);
- Clearness and comprehensibility of the material read in the publication;
- Which sections of the publication were used;
- Suggestions for improvement to any aspect of contents, usability, etc. of the publication; and
- Comparison of ease of use, etc. of the online version with the hard copy version.

The sample design and methodology also permitted us to pose key questions about **non-user** awareness of IRS electronic forms and publications, attitudes and beliefs about their use, willingness to use these forms, reasons for use and non-use, and satisfaction or dissatisfaction with the products and services. The number of relevant questions for non-users was far more limited than for users, but the information from non-users could be critical in improving the design of the form to make it more attractive and useful to taxpayers.

Methodology: Telephone survey

A national survey of taxpayers was conducted by telephone in order to generate national estimates concerning the use and perceptions of both the paper and electronic version of Publication 17, and the IRS website from which the electronic version can be obtained.

The sample of taxpayers for the telephone survey was obtained from a national sample of telephone households selected by random digit dialing (RDD) and telephone screening. Approximately 96% of households in the United States has one or more telephone lines and, hence, would be accessible for a telephone survey. The telephone households should include an even higher proportion of the taxpayer population, since both telephone service and taxpaying are associated with higher income levels. It is estimated that about three percent of the U.S. population live in cell phone only households, so this should not represent a significant sampling limitation for this study.

The specific method of random digit dialing (RDD) sampling used for this study was the truncated, list-assisted method (as distinct from the alternative Waksberg/Mitovsky cluster procedure). The "truncated" method includes only telephone banks (last two digits of a telephone number) that have any listed telephone numbers in the sampling frame. This procedure is estimated to exclude about four percent of telephone households whose telephone numbers are located in "empty" banks at the time of the sampling. Published research has demonstrated, however, that there is no statistically significant difference between the small out-of-frame sample and the in-frame sample in their population characteristics. Moreover, within the truncated sampling frame, a true simple random sample of telephone banks can be conducted. With random assignment of the last two digits to sampled banks, all published and unpublished numbers within those banks have equal probability of selection. The alternative sampling approach requires cluster sampling to be equally cost effective, and this cluster design effect reduces the precision of sampling estimates compared to a simple random sample.

The telephone numbers selected by random digit dialing were called to determine whether they were working residential numbers. If they were, then an adult in the household was screened to determine if there were any current taxpayers (filed federal income taxes for the previous tax year) in the household. If there was a

taxpaying unit in the household, then the most (or equally) knowledgeable person about the preparation and filing of the 2002 tax return was selected for an interview.

A total of 10 attempts were made to reach a household over a three week period. Calls were scheduled on evenings and weekends, primarily, according to a calling algorithm designed to maximize contact rates. The fifth or sixth call was made during the day, primarily to identify ineligible businesses with numbers in the residential telephone banks. If a household was reached within the 10 contact attempts, up to 10 additional attempts were made to complete an interview with an eligible respondent. Appointments or unscheduled callbacks were made, if the respondent was not available. One refusal conversion attempt was attempted for "soft" initial refusals.

The study design called for a total sample size of 2,000 completed interviews with a national sample of taxpayers in order to generate a sub-sample of taxpayers familiar with Publication 17, as well as a sub-sample of taxpayers who had visited the IRS website. The telephone survey was conducted between January 23, 2004 and February 8, 2004. A total of 2,000 interviews were completed with qualifying taxpayers. The telephone interview averaged 13.7 minutes in length.

Methodology: Internet Survey

During the development of the national telephone interview, an extended qualitative research phase was conducted to understand taxpayer awareness, use and attitudes toward IRS forms, publications and services. One phase of the qualitative research involved nine focused group discussions with taxpayers concerning their use of Publication 17 and other hardcopy forms and publications from the IRS; their awareness and use of the IRS website; and their awareness, use and attitudes toward Publication 17 in its electronic format on the IRS website. Another phase of the qualitative research involved web-based interviews with a national Internet sample of taxpayers.

Web-based interviews with a national Internet panel was selected as a fast and cost-effective way to generate large-scale, qualitative research concerning the use and perceptions of electronic version of Publication 17 on the IRS website among taxpayers who currently use the Internet. An Internet Panel provides a nationally dispersed sample of Internet users who have been intercepted at web-portals and agreed to re-contact for Internet surveys. Hence, it provides a cost effective method for identifying a large convenience sample of regular Internet users, which is the population from which both potential and current users of electronic Publication 17 are drawn. As appropriate with qualitative research, our primary goal was to obtain a reasonably large, geographic stratified sample of the target population for purposes of exploratory, qualitative research on the topic, rather than population estimates.

Internet panels provide an efficient sampling frame for qualitative research among the population of Internet users. Since these panels are obtained by intercepts at Internet portals, they are more likely to be representative of active users of the Internet than all adults with Internet access. Hence, while this type of sample does not permit sample estimates that can be generalized within statistical limits to the total population of taxpayers, they provide cost-efficient samples of active Internet users who are most likely to have used the IRS website and its electronic publications.

The sample of Internet users was provided by Survey Sampling Incorporated (SSI). Their SurveySpot Internet Sampling Panel is recruited by random web intercepts and agrees to accept invitations to surveys. This type of sample is recognized as an "opt in" sample that has agreed to receive Internet survey invitations, so e-mail invitations to them do not violate the increasingly tough "anti-spam" rules of e-commerce.

The panel was restricted to the United States and geographically stratified by Internet users. This unclustered, geographically stratified sample of adult Internet users should provide a sample of active Internet users who are willing to participate in Internet surveys, i.e., share their attitudes, experiences and preferences with researchers in this setting.

The methodology for qualitative interviews using an Internet panel is similar to other forms of in-depth interviews. A sample of panel members is invited by e-mail to participate in the survey. The e-mail invitation includes the purpose of the survey, sponsor, confidential and voluntary nature of the data collection, along with the web address of the survey and a personal identification number to access the survey.

The questionnaire for the web survey was estimated to take about fifteen to twenty minutes to complete. The response rate to web surveys of this length with a single invitation was expected to be about 10%. The non-participant bias in most Internet surveys is generally associated with interest in the subject matter of the survey since the e-mail invitation includes a description of the nature and purpose of the survey and each respondent decides whether to respond to the particular survey invitation. This non-participant bias could cause serious problems in estimates of the prevalence of a particular behavior in a target population, but it actually increases the efficiency of recruiting a sample of persons who have experience and opinion about a topic for qualitative research.

A total of 9,000 e-mail invitations to participate in the web-survey were sent out to three replicate samples of 3,000 members of the Internet panel. Only one e-mail invitation was sent out per subject. The incentive to participate in the survey was an opportunity to be included in a monthly drawing for prizes conducted by SSI for its panelists. A total of 947 interviews were completed by qualifying respondents to the Internet taxpayer survey. Those who qualified for a longer web interview based on their use of Publication 17, the IRS website or electronic Publication 17 were offered an additional ten dollars to complete those additional sections of the interview.

E-Mail Invitation

Dear Panelist:

We are conducting a survey for the Internal Revenue Service (IRS) concerning taxpayer satisfaction with forms, publications and services to assist customers in understanding, preparing and filing their federal income tax. We are particularly interested in the opinions of Internet users to help the IRS improve their products and services for the Web.

The interview averages less than fifteen minutes to complete. The survey is being conducted for the IRS by an independent survey research organization, SRBI, who has pledged to protect the confidentiality of your responses. If you complete the survey, your name will be entered by SSI into the monthly drawing for one of over 100 prizes worth a total of $10,000. In addition, your name will be entered in a second smaller drawing for one of five $100 prizes.

You can access the survey by clicking on the link to the IRS Survey site below or you can copy the URL into your browser. To begin the web survey, you will be asked to enter your Personal Identification Number (PIN) shown below. This is to make sure that only those who were invited participate in the survey.

If you have any questions or encounter any technical problems with the Internet survey, you can call our toll free number (1-800-659-5432) and ask for the IRS Survey Coordinator or send us an e-mail at irssurveys@srbi.com.

Thank you in advance for your participation.

John M. Boyle, Ph.D.
Senior Partner
SRBI

Respondent Characteristics

As expected, the Internet sample showed a notable bias against older respondents compared to the telephone survey. The age distribution of the national probability sample of taxpayers from the RDD telephone survey and the Internet sample was similar for the 18-34 year olds (22%-23%) and the 35-44 year olds (23%-24%). However, the telephone sample had fewer respondents among the 45-54 year olds (22%-27%) and the 55-64 year old cohort (15%-18%). By contrast 14% of the sample in the telephone survey was aged 65 and older, compared to only 6% in the Internet survey. (Figure 1)

Remarkably, the proportion of women (52%) and men (48%) in the telephone survey matches the gender distribution among the total population of adult Americans. In most national telephone surveys, a random or systematic selection procedure (e.g., most recent birthday) for eligible respondent within the household usually produces a 60%/40% female to male ratio in the completed sample. The selection procedure for the IRS telephone survey, by contrast, was to ask for the person most familiar with the filing and preparation of taxes. This is likely to have increased the proportion of male respondents in the telephone sample. The Internet survey however, could not ask for the person who was most familiar with tax filing and preparation; it could only disqualify potential respondents who were not familiar enough to answer questions about their 2002 tax return. Consequently, the proportion of men in the Internet survey was 40% compared to 48% in the telephone survey (Figure 2).

There was no difference between the telephone survey (13%) and the Internet survey (13%) in the proportion of respondents who had a post-graduate degree. However, significantly fewer Internet survey participants (16%) had a high school degree or less, compared to telephone survey participants (34%). Consequently, there were a larger proportion of survey respondents with some college or a four-year college degree in the Internet survey (70%) than in the telephone survey (52%). (Figure 3)

The Internet survey under-represents lower income taxpayers compared to the telephone survey. Among those reporting household income, a majority (58%) of survey participants in the national telephone survey of taxpayers reported a household income of less than $50,000 a year compared to only 42% of the Internet survey participants. Conversely, significantly more Internet survey respondents (58%) than telephone survey participants (43%) reported household incomes over $50,000. In fact, there is an 11% difference in the percent of Internet survey respondents who reported incomes more than $75,000 (33%) and the telephone survey participants (22%) who reported the same. (Figure 4) Twelve percent of telephone survey respondents and seven percent of Internet survey respondents were not sure or refused to disclose their household incomes.

As noted earlier, the respondent selection process was different for the telephone survey and the Internet panel. The Internet panel was limited to the person in the household who had agreed to be a member of the Internet panel, and who subsequently received an e-mail invitation to participate in the survey. Those who were not sufficiently familiar with their 2002 tax return to answer questions were screened out of the Internet survey. But there was no requirement that the Internet survey participant be the most knowledgeable or even equally knowledgeable about their tax filing and preparation.

By contrast, the telephone survey asked whether the initial respondent in the household was the adult in the household who was most familiar with the preparation and filing of their 2002 federal income tax return. If the respondent was the most familiar or equally familiar with another household member, he or she would be the designated respondent for the survey. However, if the initial contact did not feel that he or she was the most familiar, then the interviewer asked for the household member most familiar with the taxes.

As a result, nine out of ten respondents in the telephone survey (91%) were the person most familiar with preparation and filing of the 2002 federal income tax. By contrast, only two thirds (66%) of respondents in the Internet survey were the most familiar. More participants in the Internet survey described themselves as equally familiar (21%) about tax preparation and filing than in the telephone survey (9%). The Internet survey permitted participation if the respondent was sufficiently familiar to answer questions about their 2002 federal income or said they were not sure if they were familiar enough, even though they did not feel they were the person in the household most familiar with the taxes. As a result, about one in eight (13%) participants in the Internet survey were not the person in the household who was most familiar or equally familiar with preparation and filing of the 2002 federal income tax (Figure 5).

Making the Two Samples More Equivalent

The difference in the two selection procedures for the surveys introduces a clear bias in the familiarity of the survey respondent with tax preparation and tax filing. To correct this bias, we have restricted the comparison samples to the persons in both samples who were most familiar with the preparation and filing of their 2002 federal income tax. This reduces the sample size for the telephone survey to 1,820 and the sample size for the Internet survey to 618.

The demographic differences noted in the full telephone and Internet survey are primarily the result of differences between Internet users and the general adult population, rather than how familiar the respondents were with their income tax forms. Hence, demographic biases persist after limiting both samples to the persons most familiar with preparation and filing of income taxes. Proponents of the use of Internet surveys to generate sample estimates for the total population argue that most or all

biases in Internet surveys can be eliminated by weighting the demographics to population parameters. Hence, we can weight the Internet survey participants who are most familiar about their federal income taxes to approximately the demographic characteristics of the telephone survey participants who are most familiar with them.

The four demographic questions shared by the two surveys were age, gender, education and income. There was little or no missing data on age, gender or education in either survey. However, a significant proportion of the survey participants refused to answer the income question, and the proportion of missing values was significantly higher in the telephone survey than the Internet survey. Hence, we limited the target parameters to age, gender and education. The expected number of cases in each cell of the stratified design in the Internet sample was estimated based on the total number of most familiar completes in that sample (618) and the proportion of completes to total number of most familiar completes in the equivalent cell of the telephone survey sample. The actual number of completes in each cell of the Internet sample was divided by the expected number of completes based on the telephone survey to generate a cell weight. This cell weight was applied to the Internet respondents to correct for differences in gender, age and education for the telephone survey.

After both samples were restricted to the person in the household most familiar with 2002 federal income tax, and the Internet sample was weighted as described above, there was no significant difference between the two samples in age, gender or education. The income difference between the two samples was also reduced after sample restriction and weighting. Nonetheless, significantly fewer persons in the Internet sample (13%) than the telephone survey (21%) reported incomes under $25,000. Conversely, significantly more respondents in the Internet survey (50%) than the telephone survey (37%) reported incomes of $50,000 or higher even after weighting by gender, age and education.

Internet Access

Two thirds (67%) of the national telephone survey of taxpayers reported that they had Internet access at home. The Internet penetration among taxpayer households in the telephone survey closely matches estimates of the current prevalence of Internet connections in all U.S. households. By contrast, nearly all of those who participated in the Internet survey (97%) had Internet access at home. Even after restricting the sample to those who were most familiar with their federal income taxes and weighting by age, gender and education, this difference in home access to the Internet between the telephone (66%) and Internet sample (96%) persisted (Figure 6).

Internet Use and IRS Website

Since there is a relatively dramatic difference in Internet access between the two samples, even after weighting, we would expect to find difference in awareness and use of IRS products and services based on the web. So, we first investigated the awareness and use of the IRS website in the two (weighted) samples.

Seven out of ten taxpayers (72%) in the telephone sample said they were aware that the IRS had a website where they could get forms and publications. Taxpayers in the Internet sample were more likely (82%) to be aware of the IRS website. Only 18% of taxpayers in the Internet sample, compared to 27% in the telephone sample said they were unaware that the IRS had a website --- where they could get forms and publications (Figure 7).

Although the difference in taxpayer awareness of the IRS website was statistically significant, the differences were relatively modest. In contrast, there was a very big difference in taxpayer behaviors between the two samples. Nearly half (49%) of the Internet sample, compared to only 18% of the telephone sample of taxpayers had tried to use the IRS website during the 2002 tax season (Figure 8).

Taxpayers who had ever visited the IRS website were asked to rate the IRS website compared to other websites that they had seen. The telephone sample of IRS website users tended to rate the IRS website lower than the Internet sample of users. Less than two in five IRS website users from the telephone sample rated the IRS website as one of the best (4%) or above average (34%) compared to other websites they had seen. By contrast, 7% of Internet sample IRS website users rated it as one of the best, while another 37% rated it above average. The telephone sample of users were more likely to rate the IRS website as average (52%) than the Internet sample (47%). The telephone sample was also more likely than the Internet users (6%-3%) to say they were not sure how to rate it compared to others (Figure 9).

Taxpayers who had used the IRS website in the past were asked to rate their overall satisfaction with that website. This overall rating was done after a series of ratings of different aspects of the content, layout and features of the website. Taking all of these specific factors into account, the majority of all users said that they were at least somewhat satisfied with the IRS website. However, three quarters of users from the telephone sample (75%) compared to 83% of the Internet sample users, gave the IRS website an overall satisfied rating (4 or a 5 on a scale from 1 to 5 where 1 means very dissatisfied and 5 means very satisfied). The difference between the ratings of the telephone and Internet sample users was primarily in the middle range of neither satisfied or dissatisfied (17%-10%) rather than in the dissatisfied range (6%-5%) of scores of 1 or 2 on the five point scale (Figure 10).

The user's likelihood of using the IRS website in the future was treated as a behavioral measure of user satisfaction with the IRS website. Nearly nine out of ten past users of the IRS website said they definitely or probably would use the IRS website in the future in both the telephone and Internet samples. However, only 47% of the telephone sample users compared to 60% of Internet sample users said that they would definitely use the IRS website again. Only about one in ten past users from the telephone sample said that they probably (9%) or definitely (2%) would not use the IRS website in the future, compared to only 3% of the Internet sample of IRS website users (Figure 11).

Taxpayer Characteristics: 2002 Federal Tax Return

The majority of survey participants (56%) in the telephone survey reported that they had used a paid preparer to fill out their 2002 federal income tax return. Only 36% of the national telephone survey of taxpayers reported that they or their spouse had completed their 2002 federal tax return (or Telefile work sheet) by themselves. This pattern is reversed in the Internet panel survey, even after weighting the demographic characteristics to be the same as the telephone sample. The majority of the Internet panel of taxpayers (60%) said that they or another family member prepared their 2002 tax return themselves while only 36% said that they used a paid preparer for their 2002 federal tax return (Figure 12).

Even among those who prepared their own tax returns, there were differences in the method of preparation between the two samples. The telephone sample of self-preparers was nearly evenly split between those who used computer software to prepare that return (46%) and those who prepared their 2002 return by hand (43%). By contrast, self-preparers in the Internet panel were far more likely to use computer software to prepare their federal income tax return (58%) than do it by hand (36%). (Figure 13)

There were also differences between the two samples in how they filed their federal income tax return. The majority of taxpayers (56%) in the telephone sample reported that they filed their 2002 federal tax return by mail, while about two out of five taxpayers (39%) said that their 2002 tax return was filed electronically. More of the taxpayers in the Internet sample said that they filed their 2002 tax returns by mail (49%) than electronically (47%). (Figure 14)

All taxpayers were asked if they tried to use any of the taxpayer services offered by the IRS in the past year, either by phone, mail, e-mail and walk-in office. Most taxpayers in both the telephone and Internet samples reported no use of any of these taxpayer services offered by the IRS in the past year. However, fewer taxpayers in the Internet sample (59%) than the telephone sample (70%) report no use of any of these forms of taxpayer services in the past year. The differences between the telephone sample and Internet sample of taxpayers was very small for past year use of the IRS toll free number (14%-12%), the IRS walk-in office (4%-2%), and regular mail to the IRS (3%-4%). By contrast, the proportion of taxpayers who reported that they had tried to

use the IRS website in the past year was nearly twice as high among the Internet panel (29%) than the telephone sample (15%). The proportion of taxpayers using e-mail to the IRS was also higher in the Internet (4%) than telephone sample (1%), although the proportions are low in both groups (Figure 15).

Forms and Publications

Taxpayers were asked to rate their satisfaction with various aspects of getting and understanding the 2002 tax forms and instructions on a scale from 1 to 5, where 1 was very dissatisfied and 5 was very satisfied. The proportion of taxpayers (both telephone and Internet) with satisfied (scores 4 or 5) ratings was highest on knowing where to find the forms and instructions. Nearly half of all taxpayers in both samples gave satisfied scores for the completeness of the instructions and knowing what is needed to complete the return. Somewhat fewer were satisfied with the ease of determining which form is needed, the ease of understanding the forms, and the ease of understanding instructions. The lowest proportion of satisfied taxpayers in both samples was associated with the ease of finding answers in publications. Among the seven attributes rated by taxpayers, the only difference between the telephone sample and the Internet sample that was statistically significant was the proportion satisfied with ease of understanding the forms between the telephone (45%) and Internet (40%) samples (Figure 16).

Using the same five point scale, and taking all of these factors into account, taxpayers were asked their overall satisfaction with the ease of getting and understanding tax forms and instructions. Despite only one significant difference between the two samples in attributes of forms and instructions, the overall satisfaction with the ease of getting and understanding forms is significantly higher among the telephone sample (50%) than the Internet sample (44%). The telephone sample was more likely to be very satisfied (22%) than the Internet sample (12%). The Internet sample is more likely than the telephone sample to be neutral (32%-26%) and dissatisfied (23%-14%) in their overall satisfaction with the ease of getting and understanding tax forms and instructions (Figure 17).

Publication 17

Publication 17, Your Federal Income Tax, is a comprehensive tax guide for individuals that contains the Service's most comprehensive discussion of the provisions of federal income tax law that affect the vast majority of individuals in the country. It is the largest publication and the most widely used of the technical information publications issued by the Service. The survey provided current estimates of taxpayer awareness and use of this primary tool for preparing federal income tax returns.

The national sample of taxpayers was asked whether they were familiar with IRS Publication 17: Your Federal Income Tax for Individuals. In order to avoid confusion, the interviewers specified that the publication included all of the general instructions for completing federal income tax forms. They further specified that the publication was about 300 pages long and was not part of the 1040 tax packet sent by the IRS. This publication has to be ordered separately.

The survey found that less than one in five taxpayers said that they were familiar with Publication 17. There was no statistically significant difference in the familiarity with Publication 17 between the telephone sample of taxpayers (17%) and the Internet sample of taxpayers (19%). This is somewhat surprising given the higher proportion of self-preparers in the Internet sample, who might be expected to be more likely to use Publication 17 (Figure 18).

Similarly, there is no statistically significant difference in the use of Publication 17, either lifetime or during the most recent tax season, between the telephone and Internet samples. Approximately one out of ten taxpayers in the telephone sample (10%) and the Internet sample (12%) reported that they have ever used Publication 17 when preparing a tax return (Figure 19). About one in twenty taxpayers in the telephone sample (5%) and Internet sample (6%) of taxpayers reported that they used Publication 17 to prepare their 2002 federal income tax return (Figure 20).

There were differences between the two samples, however, in where they obtained this publication. The Internet sample of taxpayers was more likely than the telephone sample to have obtained the tax guide from the IRS website (38%-12%) or by e-mailing the IRS (8%-1%). By contrast, the telephone sample of taxpayers was three times as likely as the Internet sample (21%-7%) to have obtained the publication from the post office. Both samples were about equally as likely to have gotten Publication 17 from the library (23% [telephone] -29% [internet]). (Figure 21)

By contrast to the earlier measure of satisfaction with the ease of getting and using IRS forms and publications, the Internet sample was more likely to give Publication 17 a higher rating than the telephone sample. Compared to other IRS publications they have seen, the majority of the Internet sample of users of Publication 17 rated it as one of the best (17%) or above average (53%). By contrast, less than half of the telephone sample of Publication 17 users rated it as one of the best (13%) or above average (33%). Relatively few of either telephone sample users (5%) or Internet sample users (3%) rated the publication as below average or one of the worst. The telephone sample users (41%) were more likely than the Internet sample users (25%) to rate the publication as about average (Figure 22).

Once again, after a rating of a series of specific features of the publication, those taxpayers who had used Publication 17 in the past were asked to rate their overall satisfaction with the publication. The vast majority of both samples gave the publication a positive rating. However, the Internet sample of users was more likely than the telephone sample to give the publication positive ratings (80%-77%) and negative

ratings (16%-7%). The telephone sample was more likely than the Internet sample to choose the neutral rating (15%-4%). (Figure 23)

Conclusions and Implications

A comparison of the sample estimates of tax preparation and tax filing behavior between a national survey of taxpayers conducted by telephone using RDD sampling and a national survey of taxpayers conducted by Internet from an Internet panel demonstrates that self-selection biases are not eliminated by sample weighting. Even after adjusting by age, gender, education and income, the sample estimates for the most central taxpayer behavior – who prepares the income tax return – is almost the exact opposite between the two modes. Many of the other measures of interest to this research differ significantly between the two modes.

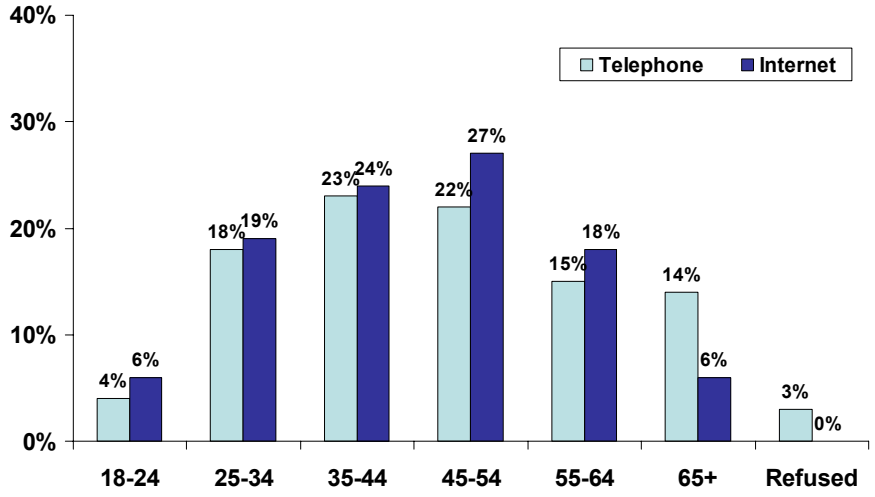
The subject matter of the survey was described in the e-mail invitation as taxpayer satisfaction with IRS forms, publications and services to assist customers in understanding, preparing and filing their federal income tax. The same introduction was used in the telephone survey. However, while telephone interviewers "push" all designated respondents to participate in the interview regardless of their interest or perceived competence in the subject matter, the Internet invitation tended to disproportionately "pull" those with interest and experience in tax preparation and filing. When the subject matter of the Internet survey is disclosed, salience of the subject matter will prompt self-selection. The incentives associated with participation in the Internet survey provide some counterbalance to salience as the primary reason to participate. Nonetheless, when incentives are relatively small and there is no pressure to participate such as repeated requests for the interview, the "pull" of subject matter will tend to over-select those for whom the topic is salient.

This tendency of the Internet survey panel to pull those with interest and relevant experience into the sample makes it an extremely valuable tool for qualitative research. The Internet Panel for the IRS survey yielded larger sub-samples of persons experienced with tax preparation and filing, and hence, the forms, instructions and documents being evaluated. The Internet survey technology also allowed us to show the respondent examples of documents within the interview or send them to website to view and use, and return. This is another extremely valuable aspect of Internet based data collection.

Nonetheless, unless the Internet panel survey incorporates a "push" component for participation that is equivalent to a well designed telephone, mail or face-to-face survey, it is unlikely to yield unbiased estimates of the population. The self-selection of the achieved sample based on the salience of the subject matter can be minimized by non-disclosure of subject matter. However, this would be a breach of any rules regarding consent in research protocol. Moreover, this would leave the rather minimal incentives as the only "pull" for participation in the survey, which would probably introduce other forms of self-selection bias in the achieved sample. Hence, the standard protocols of Internet panel surveys are unlikely to yield valid estimates of population characteristics.

FIGURES

Respondent Age



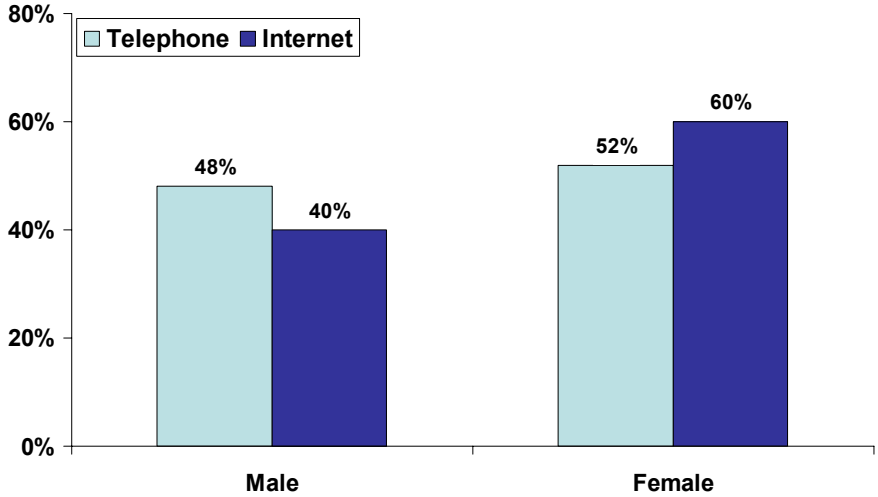
TQD1. What is your age? N=2000

IQD1. What is your age? N=947

* Difference between Telephone and Internet is significant at p<.05

1

Respondent Gender



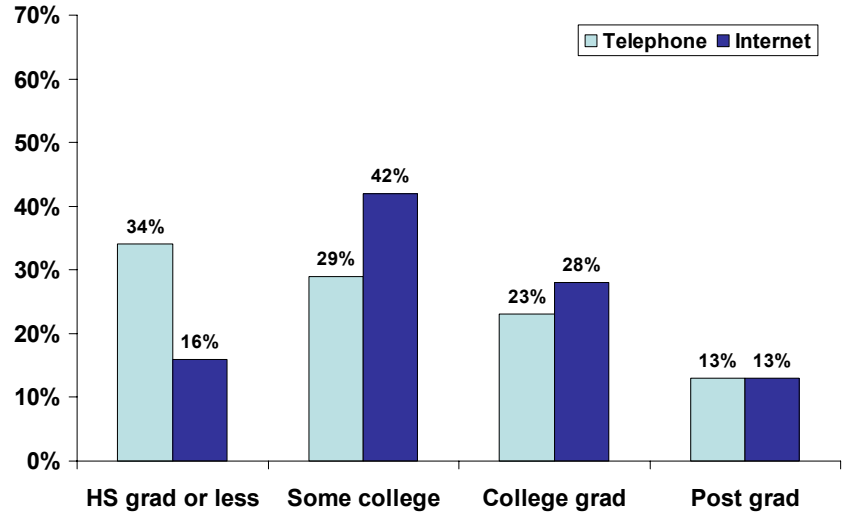
TQD10. Respondent's gender: N=2000

IQD2. Is your gender…? N=947

* Difference between Telephone and Internet is significant at p<.05

2

Respondent Education



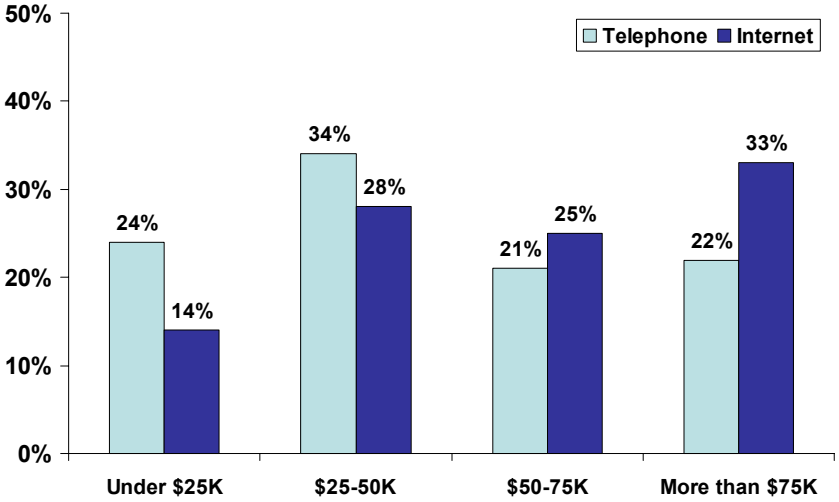
TQD4. What is the last year or grade of school you completed? N=1971

IQD3. What is the highest level of education you have completed? N=947

* Difference between Telephone and Internet is significant at p<.05 (combined: Less than HS and HS grad and excluded: Not sure/refused)

3

Household Income



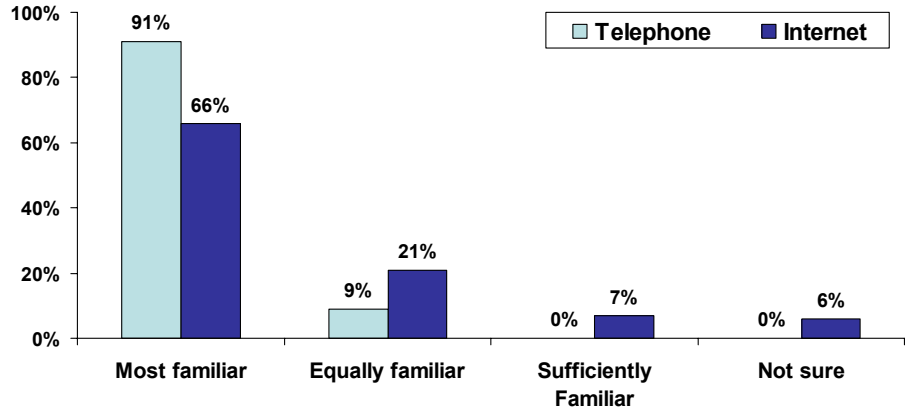
TQD7. Which of the following categories best describes your 2002 household income before taxes? N=1752

IQD4. Was your annual household income for 2002? N=878

* Difference between Telephone and Internet is significant at p<.05 (excludes: Not sure/refused)

4

Familiarity with Preparation and Filing of 2002 Federal Tax Return



TQ2. Would you say that you are the adult in your household who is most familiar with the preparation and filing of your 2002 federal income tax return? N=2000 (Tax filers)

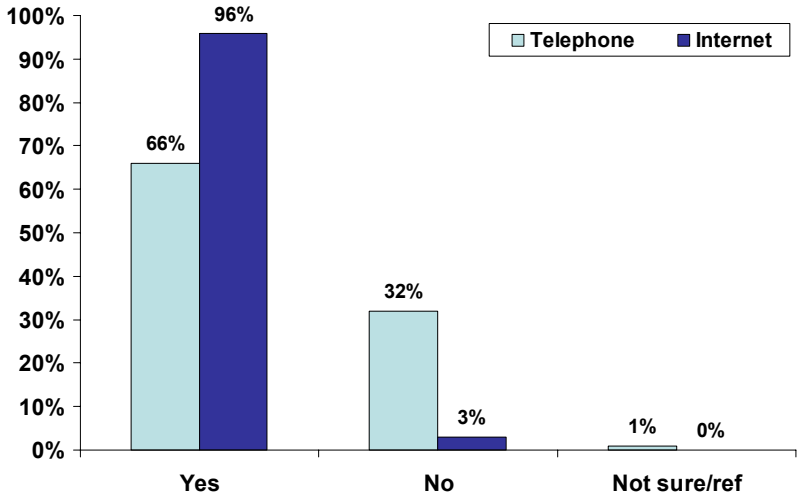
IQ2a. Would you say that you are the adult in your household who is most familiar with the preparation and filing of your 2002 federal income tax return?

IQ2b. Are you sufficiently familiar with your 2002 federal income tax return to answer some questions about it? N=932 (Tax filers in household)

* Difference between Telephone and Internet is significant at p<.05

5

Internet Access at Home



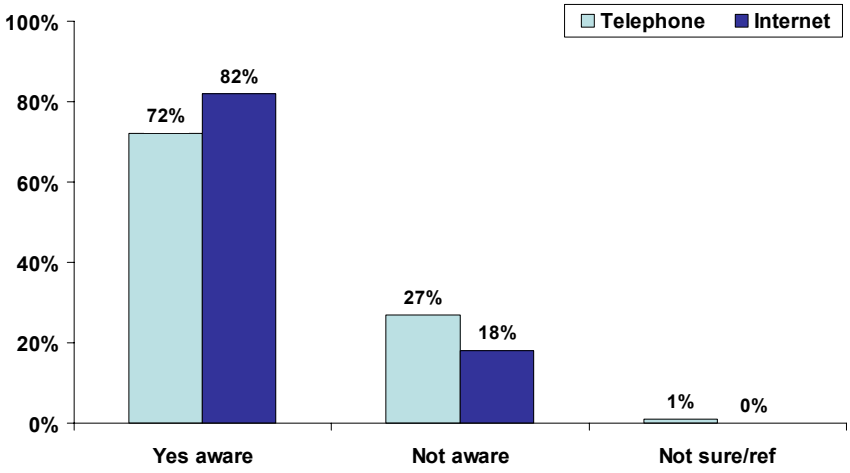
TQD9b. Do you have Internet access at home? N=1820

IQD5. Do you have Internet access at home? Unweighted N=618

* Difference between Telephone and Internet is significant at p<.05 (combined: No and Not sure)

6

Taxpayer Awareness of IRS Website



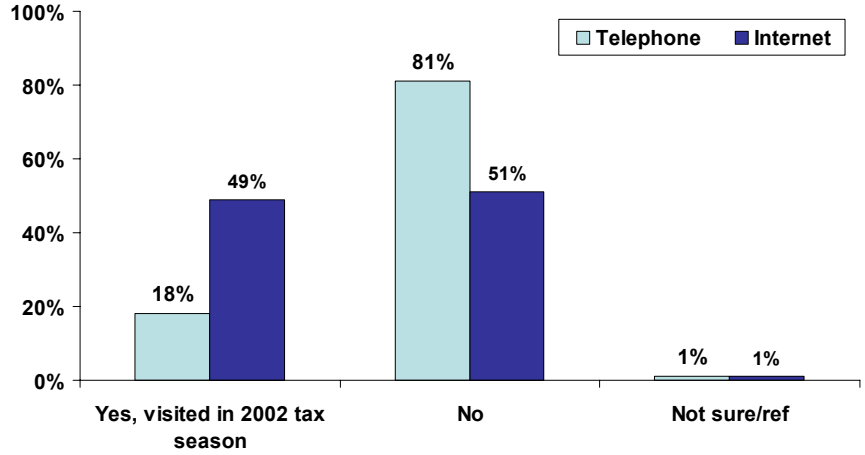
TQ23. Were you aware that the IRS had a website where you could get forms and publications? N=1820

IQ20. Were you aware that the IRS had a website where you could get forms and publications? Unweighted N=618

* Difference between Telephone and Internet is significant at p<.05 (combined: Not aware and Not sure)

7

Visited IRS Website in 2002 Tax Season



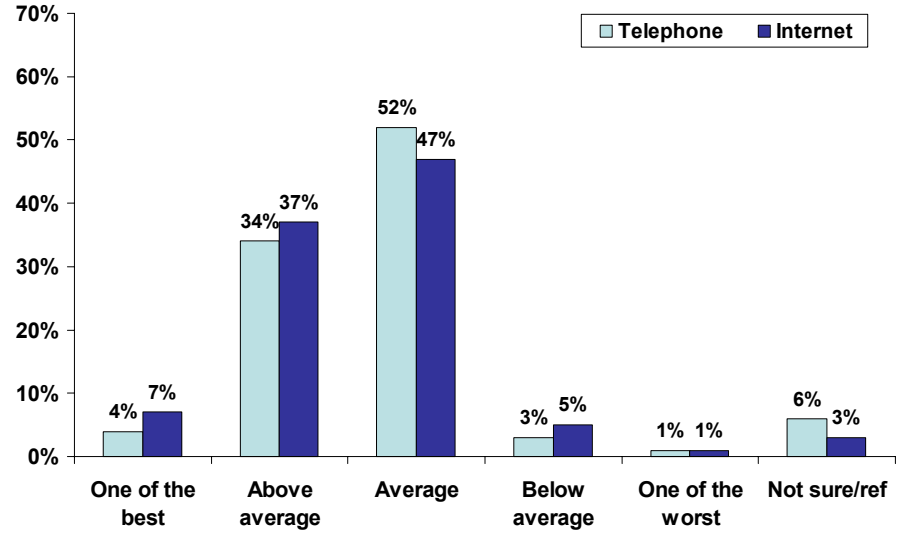
TQ24b. During the 2002 tax season, did you try to use the IRS website? N=1820

IQ21c. During the 2002 tax season, did you try to use the IRS website? Unweighted N=618

* Difference between Telephone and Internet is significant at p<.05 (combined: No and Not sure)

8

Rating of IRS Website Compared to Others

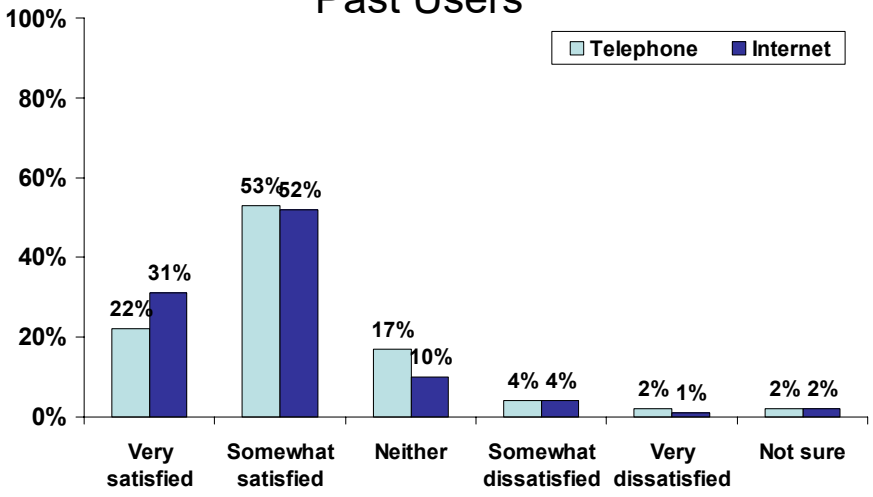


TQ25a. Compared to other websites that you have seen, would you describe this one as one of the best, above average, average, below average or one of the worst? N=573 (Have ever visited IRS website)

IQ22a. Compared to other websites that you have seen, would you describe this one as one of the best, above average, average, below average or one of the worst? Unweighted N=407 (Have ever visited IRS website)

9

Overall Satisfaction with IRS Website: Past Users



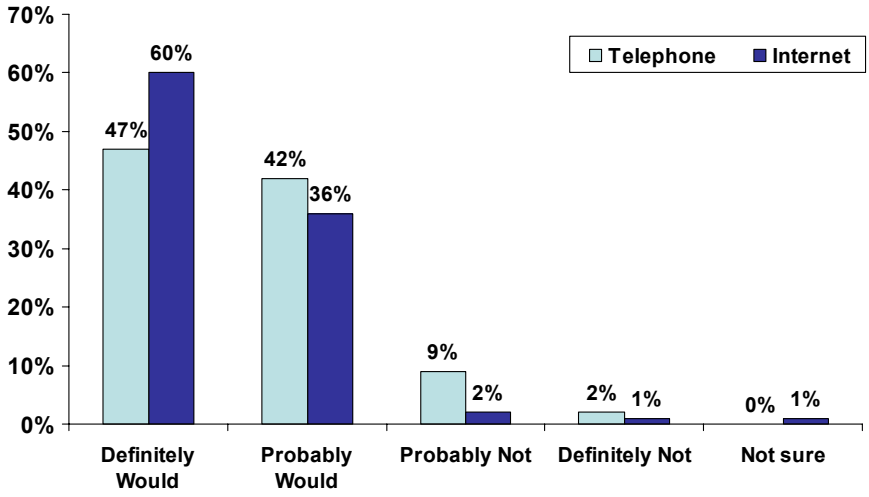
TQ29. Taking all of these factors into account, how would you rate your overall satisfaction with the IRS website? Were you… N=573 (Have ever visited IRS website)

IQ26. Taking all of these factors into account, on a scale of 1 to 5 where one means very dissatisfied and 5 means very satisfied, how would you rate your overall satisfaction with the IRS website? Unweighted N=407 (Have ever visited IRS website)

* Difference between Telephone and Internet is significant at p<.05

10

Likelihood of Future Use of IRS Website

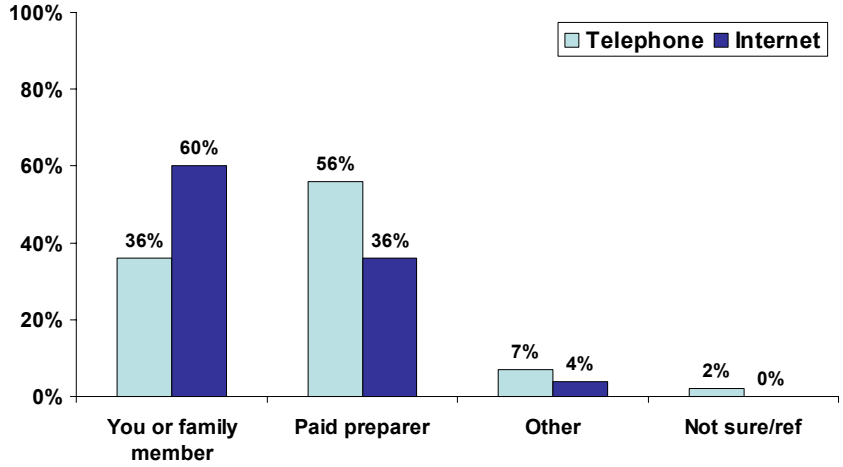


TQ30. How likely would you be to use this website in the future? N=573 (Have ever visited IRS website)
IQ27. How likely would you be to use this website in the future? Unweighted N=407 (Have ever visited IRS website)

* Difference between Telephone and Internet is significant at p<.05

11

Who Prepared 2002 Tax Form



TQ3a. Did you or your spouse complete your 2002 federal income tax return (or Telefile work sheet) or did someone else fill it out for you?

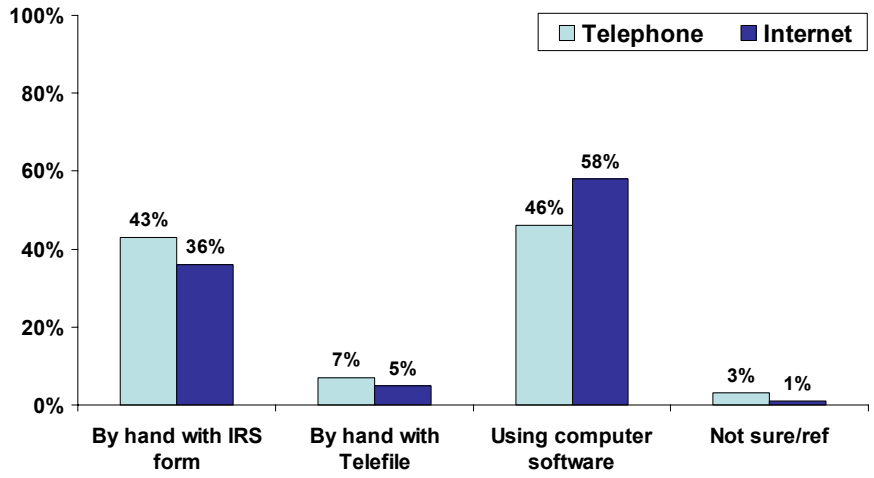
TQ3b. Who filled out your return (or Telefile work sheet) for you? N=1820

IQ4a. Did you prepare your 2002 federal income tax return yourself or did you use a paid preparer, like an accountant or tax service? Unweighted N=618 (Sufficiently familiar to answer)

* Difference between Telephone and Internet is significant at p<.05 (combined: Paid preparer, Other and Not sure)

12

How Prepared 2002 Federal Return



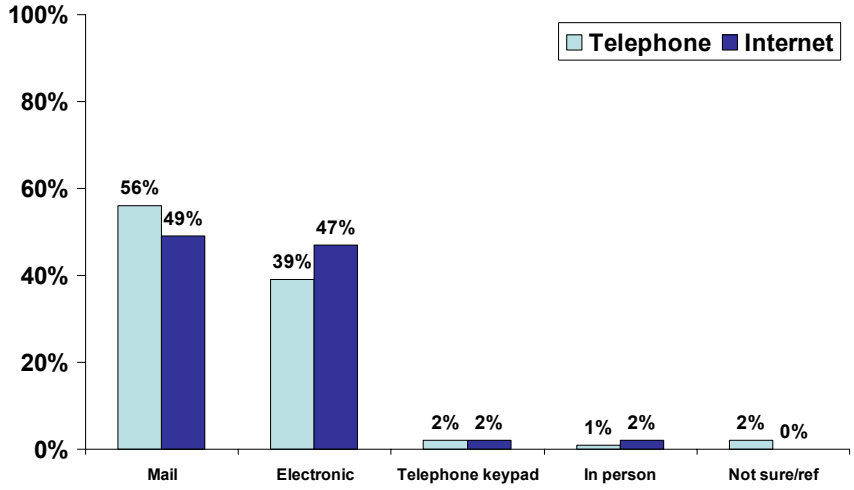
TQ4d. How did you prepare your federal income tax return? N=661 (Self or spouse prepared)

IQ4b. How did you prepare your federal income tax return? Unweighted N=421 (Self or family member prepared)

* Difference between Telephone and Internet is significant at p<.05

13

How Filed Federal Return in 2002



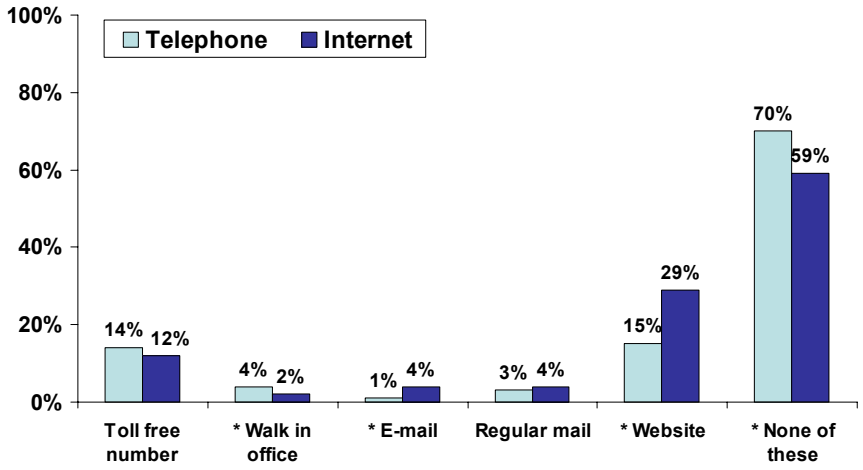
TQ9a. After your tax forms were completed, how did you file your 2002 tax return? Did you….? N=1820

IQ5a. After your forms were completed, how did you file your 2002 tax return? Did you….? Unweighted N=618 (Sufficiently familiar to answer)

* Difference between Telephone and Internet is significant at p<.05

14

IRS Services Used During 2002 Tax Season



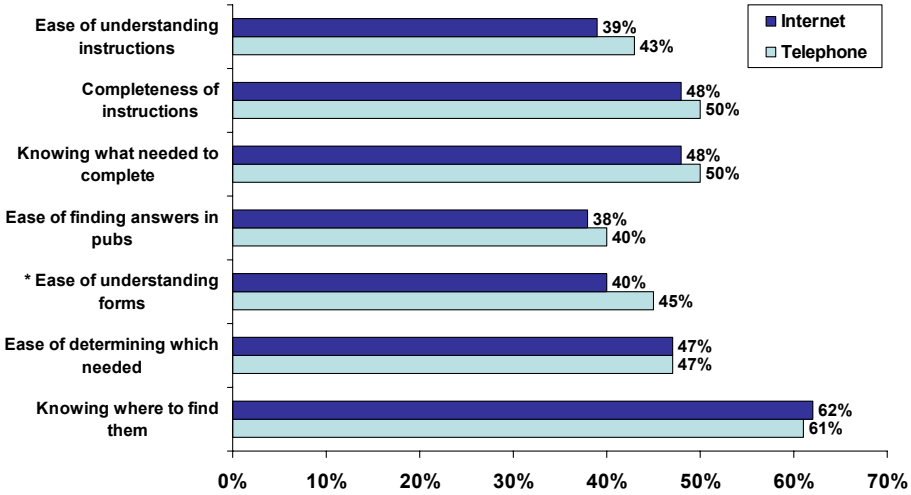
TQ10. During the 2002 tax season, did you try to use the following taxpayer services offered by the IRS? Did you try… N=1820

IQ6. During the 2002 tax season, did you try to use the following taxpayer services offered by the IRS? Unweighted N=618 (Sufficiently familiar to answer)

* Difference between Telephone and Internet is significant at p<.05

15

Very/Somewhat Satisfied with 2002 Forms and Instructions



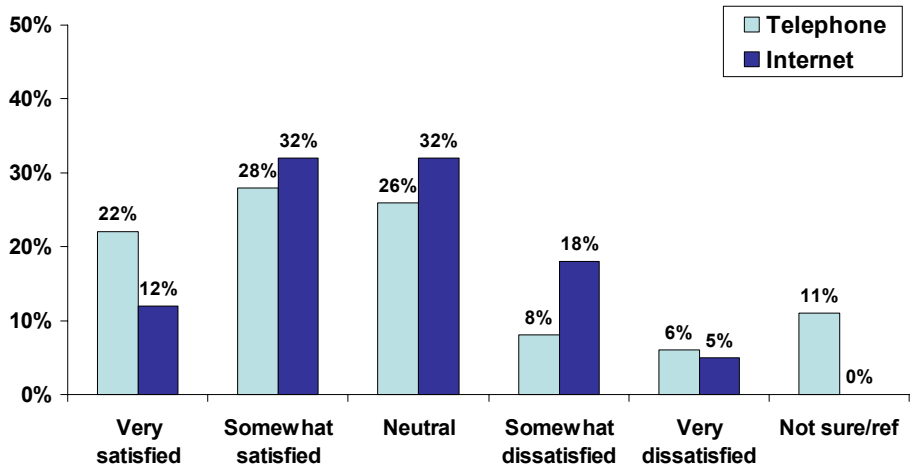
TQ13. In thinking about getting and understanding the 2002 tax forms and instructions, please rate your satisfaction with each of the following items on a scale of 1 to 5 where 1 is very dissatisfied and 5 is very satisfied. N=1820

IQ9. In thinking about getting and understanding the 2002 tax forms and instructions, lease rate your satisfaction with each of the following items on a scale of 1 to 5 here 1 is very dissatisfied and 5 is very satisfied. Unweighted N=618

* Difference between Telephone and Internet is significant at p<.05

16

Overall Satisfaction with Ease of Getting and Understanding Tax Forms and Instructions



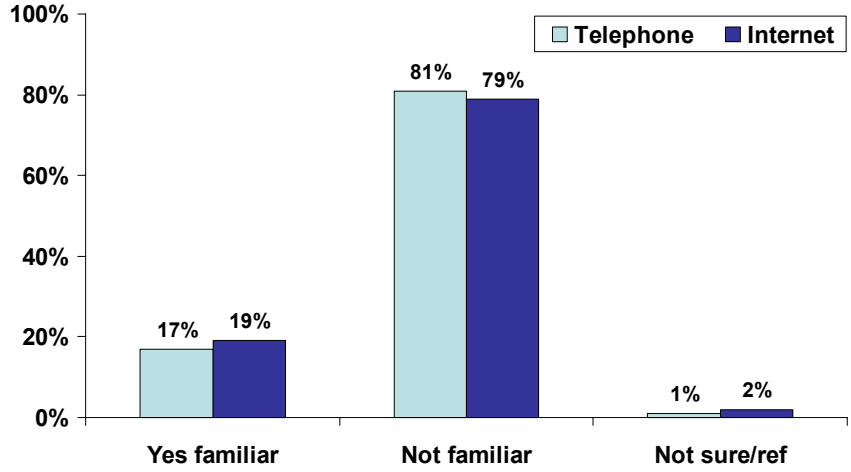
TQ14. Taking all of these factors into account, on a scale of 1 to 5 where 1 means very dissatisfied and 5 means very satisfied, please rate your satisfaction with the ease of getting and understanding tax forms and instructions? N=1820

IQ10. Taking all of these factors into account, on a scale of 1 to 5 where 1 means very dissatisfied and 5 means very satisfied, please rate your satisfaction with the ease of getting and understanding tax forms and instructions? Unweighted N=582 (Sufficiently familiar to answer)

* Difference between Telephone and Internet is significant at p<.05

17

Taxpayer Familiarity with Publication 17

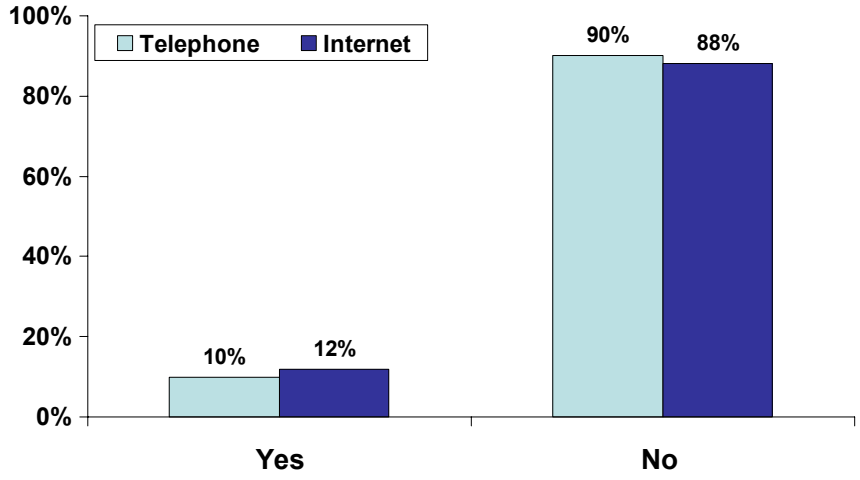


TQ15a. Are you familiar with the IRS Publication 17: Your Federal Income Tax for Individuals? It includes all of the general instructions for completing federal income tax forms. It is about 300 pages long. This is not part of the income tax you receive in the mail from the IRS. It's a separate publication that you have to order. N=1820

IQ11a. Are you familiar with the IRS Publication 17: Your Federal Income Tax for Individuals? It includes all of the general instructions for completing federal income tax forms. It is about 300 pages long. This is not part of the income tax you receive in the mail from the IRS. It's a separate publication that you have to order. Unweighted N=618

18

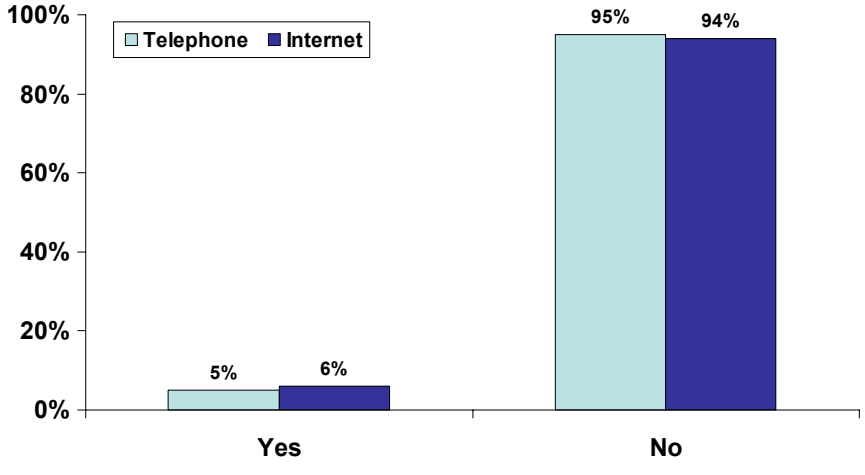
Taxpayer Use of Publication 17: Lifetime



TQ15b. Have you ever consulted this publication when preparing a tax return? N=1820
IQ11b. Have you ever consulted this publication when preparing your or someone else's tax return? Unweighted N=618

19

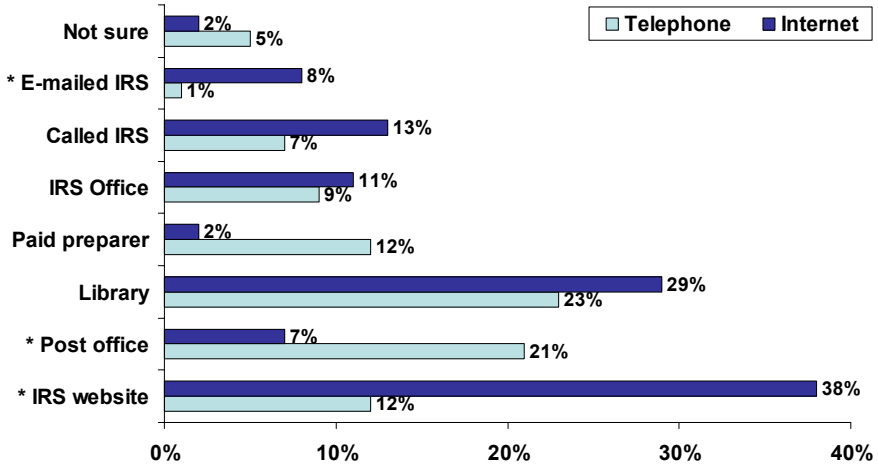
Taxpayer Use of Publication 17: 2002 Taxes



TQ5a/15c. Did you use this publication when preparing or filing your 2002 tax return? N=1820
IQ12b. Did you use this publication when preparing or filing your 2002 tax return? Unweighted N=618

20

Source of Publication 17 for 2002 Federal Tax Return



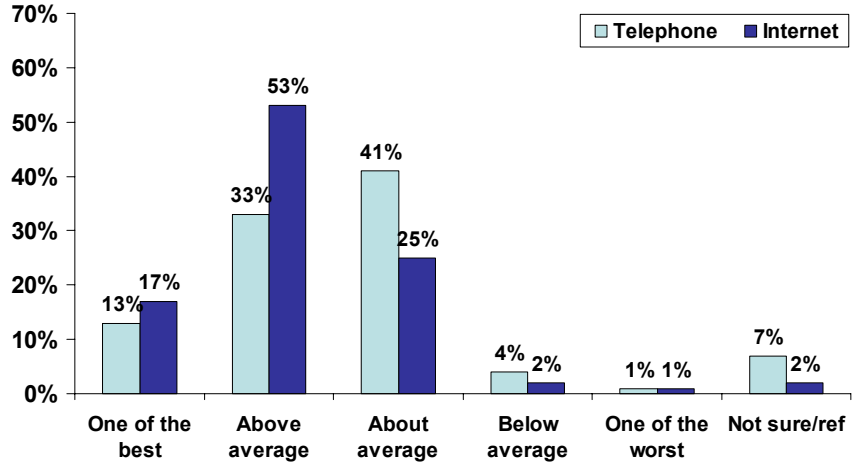
TQ5b/Q15d. Where did you get (or use) the booklet? N=94 (Used Pub 17 for 2002 tax return)

IQ12c. Where did you get (or use) the 2002 Tax Guide? Unweighted N=52 (Used Pub 17 for 2002 tax return)

* Difference between Telephone and Internet is significant at p<.05

21

Rating of Publication 17 Compared to Other IRS Publications



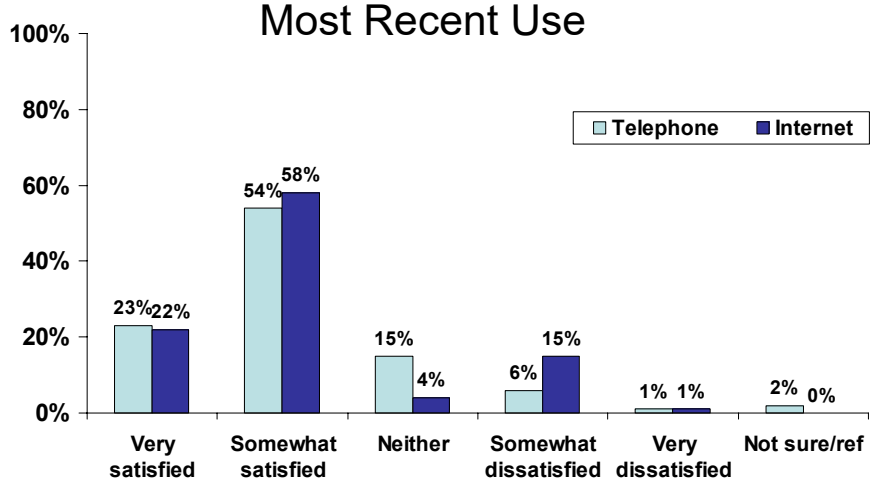
TQ17a. Compared to other IRS publications that you have seen, would you describe this as one of the best, above average, about average, below average or one of the worst? N=177 (Ever used Publication 17)

IQ14a. Compared to other IRS publications that you have seen, would you describe this as one of the best, above average, average, below average or one of the worst? Unweighted N=83 (Ever used Publication 17)

* Difference between Telephone and Internet is significant at p<.05

22

Overall Satisfaction with Publication 17: Most Recent Use



TQ21. Taking all of these factors into account, on a scale of 1 to 5 where 1 means very dissatisfied and 5 means very satisfied, how would you rate your overall satisfaction with the most recent Publication 17 that you used? N=177 (Ever used Publication 17)

IQ18. Taking all of these factors into account, on a scale of 1 to 5 where 1 means very dissatisfied and 5 means very satisfied, how would you rate your overall satisfaction with the most recent Publication 17 that you used? Unweighted N=83 (Ever used Publication 17)

* Difference between Telephone and Internet is significant at p<.05

23