ENCOURAGING RECORD USE FOR FINANCIAL ASSET QUESTIONS IN A WEB SURVEY

MICK P. COUPER*
MARY BETH OFSTEDAL
SUNGHEE LEE

We report on two experiments to encourage record use by respondents in an Internet survey. The experiments were conducted in the 2009 Health and Retirement Study (HRS) Internet Survey, administered to those in the HRS panel with Internet access, and in the 2011 HRS Internet Survey. Encouraging respondents to consult records at the relevant point in the questionnaire significantly increased reported record use (from 39 percent to 47 percent), but was insufficient to produce significant changes in the precision (amount of rounding) of the information reported. Including the encouragement in the mailed invitation to the Web survey in 2011 resulted in a lower response rate (77 percent with encouragement, 80 percent without), but increased reported record use among respondents (from 46 percent to 55 percent). In neither case was the increase in reported record use large enough to produce significant differences in the precision of the information reported between the groups with and without encouragement.

1. INTRODUCTION

It is a common belief in the survey literature that encouraging respondents in household surveys to use records will increase the quality of data obtained. For example, Laurie and Moon (2010) write, “Encouraging respondents to refer to documents is likely to lead to the collection of better quality data at little risk to respondent cooperation.” It is further argued that self-administered modes (e.g., mail and Web) impose less time pressures on respondents, and...
would facilitate the use of records to answer factual questions (e.g., Kanuk and Berenson 1975).

However, while a number of (mostly unpublished) papers have explored the levels and correlates of record use by respondents, we are aware of no studies that have attempted to increase record use experimentally. We report on two such studies, both conducted in an Internet survey among members of a large, ongoing panel study in the United States.

2. BACKGROUND

There are many factors that could affect the use of records by household members to answer survey questions. These include (1) the availability of records, (2) the ease of access (i.e., even if records were available, they may be hard to retrieve), (3) the correspondence between the information requested and the records (i.e., if consulting records would directly provide the answer or require further processing), (4) the types of information requested (i.e., some types of information may be more readily accessible in memory than others), (5) respondents’ motivation to use records, and (6) the extent to which respondents are encouraged to use records. Our focus is on the last of these factors. Further our focus is on household rather than establishment surveys. Many establishment surveys are designed to elicit answers that exist in records rather than respondents’ heads, and the record-use process in such surveys is likely to be quite different.

Most of the research on record use has focused on three key areas: (1) the proportion of respondents using records, (2) the socio-demographic correlates of record use, and (3) the data quality implications of record use. We review this limited literature briefly here. Given our research objectives, we focus primarily on the first of these in our review.

In one of the earliest reports of record use, Marquis, Moore, and Bogen (1993; see also Moore, Marquis, and Bogen 1996) reported that about 20 percent of respondents in the Survey of Income and Program Participation (SIPP) used at least one record, but provided no further details of record use. In a small-scale (sample size ≈ 100) cognitive pretest, which tested a number of different design changes, they were able to increase record use to 83 percent in wave 1 and 85 percent in wave 2 of the survey. We are unaware of any evidence that these strategies were tested or implemented on a larger scale.

More recently, Safir and Goldenberg (2008) evaluated data quality for face-to-face versus telephone respondents in the Consumer Expenditure Interview Survey (CEX). They found that record use was reported in about 39 percent of personal interview surveys and 25 percent of telephone surveys. Edgar (2010) examined interviewer reports of the extent of record use in the 2006–2008 CEX, reporting that in about 15 percent of cases, interviewers reported that households always used records, about 16 percent almost always, and 11
percent mostly, with 26 percent reporting use occasionally or almost never, and 31 percent never.

Carter (2010) examined correlates of reported utility bill use in the 2007 American Housing Survey (AHS), where use of bills is encouraged in the advance letter. Overall, he found that about 40 percent of responding households had at least one bill available upon request and 36.5 percent had all four bills. He also found higher rates of record use for personal visit than telephone interviews (OR = 1.41).

Laurie and Moon (2010) reported that between 31 percent and 43 percent of respondents consulted pay slips across 17 waves of the British Household Panel Study (BHPS).

The National Immunization Study (NIS), which is conducted by telephone, asks parents of children 19–55 months of age whether they have a written record of the vaccination history available. If so, they are asked to read information directly from the record. In an analysis of NIS data from 2004–2006, McElligott and Darden (2010) found that 41 percent of parents reported having the records available. In a similar analysis of data from the 2008 NIS-Teen survey (focusing on adolescents age 13–17), Dorell, Jain, and Yankey (2011) found that 21.3 percent of parents used immunization cards to respond to the telephone survey questions.

Kashihara and Wobus (2007) examined the accuracy of records used in the household component of the Medical Expenditures Panel Survey (MEPS-HC), in which a calendar was provided to respondents to record expenditures between waves. For out of pocket expenditures, about 60 percent reported using memory to answer, while a further 20 percent reported using the calendar, meaning that about 20 percent of respondents used one or more records. For private insurance expenditures, about 56 percent relied on some form of record, suggesting that record use differs by type of expenditure.

O’Brien (2010) reported that of about 4,400 households interviewed in the 2005 Residential Electricity Consumption Survey (RECS), 2,794 said they had an electricity bill on hand, and 1,525 said they had a gas bill on hand, but only 2,014 (about half) had at least one bill scanned by the interviewer during the interview.

Geisen et al. (2012) and Hendershot et al. (2012) report on a small (n = 115) study of volunteers with the goal of exploring the availability and feasibility of using records in the context of the CEX. At the first interview, respondents were encouraged to retain receipts and give these to the interviewer at the second interview (about 1 week later). Records were available for about one-third of expenditures, but respondent-level compliance was not reported. We note that there is likely a difference between asking respondents to provide records and asking them to consult records to formulate their answers (as suggested by O’Brien’s findings).

These studies suggest a wide range in reported record use, but in most cases a minority of respondents used records. Turning briefly to the effects of record use
on data quality and interview length (i.e., the benefits and costs of record use), Safir and Goldenberg (2008) noted that record use was associated with higher total expenditures reported, but cautioned that this could be because households with higher expenditures may be more likely to use records. In contrast, Edgar (2010) found that record use was associated with longer interviews, a higher number of reported expenditures, and higher reported expenditure amounts, but was not consistently related to higher data quality. However, in an earlier analysis of CEX data, Edgar and Gonzalez (2009) found that reported record use was associated with lower odds of needing editing, controlling for mode of interview. Contrary to Edgar (2010), Laurie and Moon (2010) found no evidence of increased interview length for those who used records.

McElligott and Darden (2010) found a significantly higher odds of the proportion up-to-date with vaccinations when records were used than when not (OR = 1.62), controlling for socio-demographic variables. Kashihara and Wobus (2006) found that record use increased accuracy over relying on memory alone. In models controlling for a variety of socio-demographic characteristics (but not for respondent reluctance or motivation), odds ratios ranged from 1.12 to 1.24 for use of records over relying on memory for out-of-pocket expenditures, while for private insurance expenditures they ranged from 1.54 to 2.90.

The assumption – whether implicit or explicit – behind these studies is that having respondents use records is a good thing, but none address the question of whether and how to encourage respondents to consult records in answering these surveys. Further, while there is some evidence that those who use records provide data of better quality, there are mixed results on whether using records increases the length of the interview (i.e., potentially increasing burden for respondents). This research note aims to address this gap in the literature, specifically with regard to record use in a self-administered Web survey. Specifically, we examine whether encouragement to use records results in increased record use.

We conducted two experiments in the context of the Health and Retirement Study (HRS; see http://hrsonline.isr.umich.edu/). HRS is a panel study of persons age 50 and older in the United States. Panel members are interviewed in person or by telephone every two years. Panel members with Internet access are invited to complete a Web survey in the “off-years” between regular waves of data collection. The first experiment was conducted in the 2009 HRS Internet survey and the second in the 2011 HRS Internet survey. We describe the design of each study in turn, along with the findings, before offering an overall summary of the results.

3. STUDY 1: 2009 HRS INTERNET

3.1 Study Design

The sample for the 2009 Internet survey was drawn from respondents who reported Internet access in the HRS 2008 Core survey, plus those who did not
respond to the 2008 Core survey but had been selected for prior Internet surveys. Roughly 20 percent of the eligible pool was reserved for a control group, and were not invited to complete the Internet survey. Mailed invitations were sent to 5,742 HRS respondents, providing the URL and login information. The letter included a $25 check for completion of the survey. The field period was from March to August 2009. A total of 4,433 respondents started the survey, and 4,371 completed it, for a simple response rate (AAPOR RR1) of 76.1 percent. Breakoff rates did not differ by experimental group, so we restrict our analyses to the 4,371 respondents who completed the survey. The Internet survey took an average of 52.8 minutes to complete.

The experiment on encouraging record use focused on a series of questions on financial assets, asked towards the end of the Internet survey. At the beginning of this section, respondents were randomized to one of two groups. The treatment or “encouragement” group received the following introduction:

“In the questions that come next we will be asking about your finances. We encourage you to use statements or records that you have available to help answer some of the questions. Records that might be helpful include bank, retirement, and investment account statements, personal finance software, and other information about your finances. When answering the questions, we would like good estimates but not necessarily exact dollar amounts if those are not readily available.

The current economic downturn differs from the past because of the large number of families who have some of their savings or retirement plans invested in the stock market. We would like to know more about the ways in which the stock market problems may have affected your life.”

The control or “no encouragement” group received only the second paragraph of the above introduction.

Approximately half of the sample was randomized to each condition. Both groups were asked at the end of the questionnaire: “What sources of information, if any, did you use to assist you in answering these questions about your finances?” The following check-all response were included: investment account statements; bank statements; check registers; personal finance software; tax returns; I asked someone else; other sources; and I did not use any records. We coded all responses except the last as self-reported use of records.

3.2 Analysis and Results

All analyses reported here are based on unweighted data. Reflecting the weights and complex sample design from the prior interviewer-administered
waves of the HRS does not substantially affect the results or conclusions. We address two research questions: (1) does encouragement increase reported record use, and (2) does this result in increased precision of asset values? In lieu of a measure of accuracy of reported assets, we explore whether respondents reported any assets or not and, if so, the precision with which they reported the asset values, as indirect indicators of data quality.

We find that those encouraged to use records did indeed report doing so at higher rates than those not encouraged (46.7 percent vs. 38.7 percent, \(\chi^2 = 28.55, \text{d.f.} = 1, p < .0001\)). However, this difference is not large (8 percentage points) and in both groups less than half the respondents reported consulting records to answer these questions.

The next question is whether this change in the proportion using records is sufficiently large to produce changes in the precision of asset measures. To examine this, we examine questions about five types of assets: (1) 401(k) or other employer-based retirement savings, (2) IRA or Keogh, (3) trust funds, (4) mutual funds, and (5) shares of individual companies. For each of these assets, respondents were first asked if they owned the asset and, if yes, what the total value of the asset was. If respondents did not provide an answer to the open field for the dollar amount, they were presented with a range of values, which included a “don’t know” option. Using these responses, we created a measure of precision along the lines first used by De Leeuw (1992). Note that this is not independent of the size of the asset – one cannot round to the nearest $100,000 if the value of the asset is less than $10,000 – but in the absence of external validation data, this serves as an approximation of precision. Also note that the introduction specifically mentions “good estimates but not necessarily exact dollar amounts.” The distributions of this measure for the five assets, by the encouragement manipulation, are shown in Table 1.

As can be seen in Table 1, there are no significant differences in the level of precision between the experimental groups across the five assets. Examination of individual outcomes (e.g., DK responses) reveals no particular pattern across the five questions. An analysis (not shown) of the precision of responses provided by those who reported using records versus those who did not report using records, showed significant differences across all five asset types. That is, those who use records are more likely to provide substantive responses (as opposed to missing data, DK responses, or a range) and to do so with greater precision than those not using records. This suggests that those who use financial records are different from those who do not (consistent with the earlier studies cited above). However, the 8 percentage point increase in record use achieved with the encouraging statement at the start of the section is not sufficient to produce meaningful differences in the quality of answers provided to these five topics over those not so encouraged.

One post hoc hypothesis for the relatively modest increase in the proportion using records and the consequent lack of effect on precision of the data
provided is that the encouragement to consult financial records may come too late (at the start of the relevant section). Respondents may be reluctant to pause in their completion of the survey to access such records. If this is the case, then encouraging respondents to assemble relevant records prior to the start of the survey (i.e., in the invitation letter) may have a larger effect. This led to the design of the second experiment.

4. STUDY 2: 2011 HRS INTERNET

4.1 Study Design

As with the 2009 HRS Internet survey, a subsample of HRS panel members was invited to participate in the 2011 Internet survey. The sample was drawn from respondents who completed their core HRS interview on or before January 31, 2011 and reported Internet access. The sample included respondents who were in the 2009 Internet sample and who still had Internet access, along with an 80 percent random subsample of new cohort respondents who
had Internet access and an 80 percent random subsample of panel respondents who reported Internet access in 2010 but not in the prior wave (newly acquired Internet access). The remaining 20 percent of the latter two groups were assigned to the control group, and not invited to the Internet survey.

Mailed invitations were sent to 5,706 HRS respondents, providing the URL and login information. The letter included a $25 check for completion of the survey. The field period was from May 2011 through September 2011. A total of 4,590 respondents started the survey and 4,503 completed it, for a simple response rate (AAPOR RR1) of 78.9 percent. The Internet survey took an average of 52.6 minutes to complete.

As mentioned above, the experimental manipulation was done in the invitation letter. The invitation letter sent to the encouragement group including the following highlighted paragraph:

<table>
<thead>
<tr>
<th>You may want to use some reference materials.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Some of the questions in this year’s internet survey are about family finances. We encourage you to use statements or records that are readily available to you when answering these types of questions. Records that might be helpful include bank, retirement, and investment account statements, personal finance software, and other information about your finances.</td>
</tr>
</tbody>
</table>

Those in the control (no encouragement) condition received no such statement. Given a concern that adding such a paragraph to the invitation letter might increase nonresponse, about one fourth of the sample was allocated to the encouragement condition, with the balance getting no such encouragement.

Because the experimental manipulation occurred at the time of invitation, the first question is whether the paragraph encouraging record use in the invitation letter decreased response to the survey. While partials (breakoffs) did not differ by experimental group (1.6 percent for encouragement, 1.5 percent for no encouragement), the completion rate did: 76.7 percent of those receiving the encouragement completed the survey, compared to 79.7 percent of those not receiving the encouragement ($\chi^2 [\text{d.f.} = 1] = 6.03, p = 0.014$).

The next question is whether the encouragement increased the reported use of records among those who completed the survey. The question about record use was the same as in the first study, and we coded self-reported record use in the same way. Among those who completed the survey, those encouraged to use records did so at a significantly higher rate than those not so encouraged (55.4 percent versus 45.6 percent; $\chi^2 [\text{d.f.} = 1] = 31.51, p < .0001$). Overall, more respondents reported using records in 2011 (48.0 percent) than in 2009 (42.7 percent), and the effect of the encouragement seems to be slightly larger (9.8 versus 8.0 percentage points), but this increase must be weighed against the higher unit nonresponse among those receiving the encouragement in the invitation letter.
Finally, does this increase in the proportion reporting using records result in improved precision of reported asset amounts? The asset section was asked of a random subset of respondents to the survey in 2011. Further, only two asset questions were asked, one about all retirement accounts (including IRA, 401(k), and Keogh accounts), the other about stocks and stock mutual funds. The questions differed from the 2009 survey in another way – the follow-up range items did not include a “don’t know” option, but respondents could still leave this item unanswered if they wished.

Table 2 shows the distributions of the precision measure for these two assets. Again, we see no significant effect on precision of record-use encouragement. As was found for 2009, those who reported using records provided responses with greater precision than those who did not (analyses not shown).

5. COMBINED ANALYSES

Given that the HRS is a panel study, and many respondents were invited to participate in both waves of the Internet survey, we can examine consistency of behavior across waves and possible carryover effects of the experimental manipulations (which were randomized independently in the two waves). Several pertinent questions could be posed with such a merged dataset, including:

(1) Does encouragement in 2009 produce higher nonresponse in 2011? Does this vary by the 2011 manipulation?

(2) Does encouragement in 2009 increase record use in 2011?

(3) How consistent is reported record use in the two waves?

Table 2. Precision of Reported Asset Values by Encouragement Manipulation, Study 2

<table>
<thead>
<tr>
<th>Type of Asset</th>
<th>Level of Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Missing</td>
</tr>
<tr>
<td>Retirement accounts</td>
<td>(1568)</td>
</tr>
<tr>
<td>Encouraged</td>
<td>(397)</td>
</tr>
<tr>
<td>Not encouraged</td>
<td>(1171)</td>
</tr>
<tr>
<td>Test of difference</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\chi^2 = 0.068$, d.f. = 4, $p = 0.99$</td>
</tr>
<tr>
<td>Stocks or mutual funds</td>
<td>(887)</td>
</tr>
<tr>
<td>Encouraged</td>
<td>(228)</td>
</tr>
<tr>
<td>Not encouraged</td>
<td>(659)</td>
</tr>
<tr>
<td>Test of difference</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\chi^2 = 6.47$, d.f. = 4, $p = 0.17$</td>
</tr>
</tbody>
</table>
We briefly address each of these in turn. The first question relates to whether there is a lagged effect of encouragement on later survey participation. Those encouraged to use records in 2009 may have completed the survey but may be less willing to complete another Internet survey, particularly if they receive the encouragement again. We see no evidence of this: 90.7 percent of those in the encouragement group and 91.8 percent of those in the no encouragement group in 2009 completed the 2011 survey ($\chi^2 = 2.28$, d.f. = 1, $p = 0.32$). We find no further differential effect by the 2011 manipulation. That is, those who were encouraged in the advance letter in 2011 and who also received encouragement to use records in 2009, did not respond at a lower rate than those not encouraged in 2009.

Second, is there a carryover effect, such that those who are encouraged twice use records at a greater rate in 2011 than those encouraged only once, or not at all? Again, we see no evidence of this. Combining the two encouragement manipulations, we find that, among those who responded to both waves, 51.8 percent of those encouraged in both waves reported using records, compared with 51.4 percent of those encouraged only in 2011, while 43.1 percent of those only encouraged in 2009 used records in 2011, compared to 45.6 percent of those not encouraged to use records in either wave. The two-year gap between waves of the Internet survey likely accounts for this lack of effect. That is, the 2009 and 2011 experiments appear to have independent effects on reported record use.

Finally, how consistent is record use across waves? Here we find a strong association ($\phi = 0.33$), with 65 percent of those reporting using records in 2009 also reporting doing so in 2011. Two-thirds of respondents (66.6 percent) were consistent in their reported record use (yes-yes or no-no), while the balance used records in only one wave (yes-no or no-yes). These combined results again suggest that record use has more to do with the type of respondent than with the effect of the encouragement.

6. DISCUSSION

Across two experiments in the same population and focusing on the same types of questions and records, we find that efforts to encourage respondents to use records to answer questions about financial assets result in relatively small but statistically significant gains in the number reporting doing so. Including such encouragement in the invitation letter increases the rate of nonresponse, but also increases the rate of record use. That is, providing sufficient lead time to assemble the necessary records may increase the use of such records, but also alerts sample persons to the upcoming request and may discourage them from participating. While those who report using records provide data of greater precision than those who report not using records, the magnitude of the increase in reported record use brought about by the experimental
manipulations is not sufficiently large to detect reliable improvements in the overall precision of estimates for these questions.

The results of these two experiments support the earlier research that respondents who use records provide answers of greater precision (higher quality) than those who don’t use records. But the reason for this may be that they are more motivated respondents. Efforts to encourage respondents who might otherwise not use records to do so, may not be that successful – unless we can convince respondents that consulting records would be of benefit to them. It seems that there is little harm in encouraging respondents to use records at the time they are needed (as we did in the 2009 study), but research is needed on ways to increase the number of respondents using records.

References


De Leeuw, E. D. (1992), Data Quality in Mail, Telephone and Face to Face Surveys, Amsterd: TT-Publikaties.


