Mode effects in mixed-mode economic surveys: Insights from a randomized experiment

Joanne W. Hsu and Brooke H. McFall

Mode effects in mixed-mode economic surveys: Insights from a randomized experiment

Joanne W. Hsu

Federal Reserve Board of Governors

Brooke Helppie McFall

University of Michigan

February 5, 2015

Abstract

Web-based surveys have become increasingly common in economic, marketing, and other social science research. However, questions exist about the comparability of data gathered using a web interview and data gathered using more traditional survey modes, particularly for surveys on household economic behavior. Differences between data from different survey modes may arise through two different mechanisms: sample selectivity due to (lack of) web access and mode effects. This study leverages the randomized experimental design of the mixed-mode Cognitive Economics Study to examine mode effects separately from sample selectivity issues. In particular, we examine differences in survey response rates, item nonresponse, and data quality due to mode effects. Our results indicate that, in contrast to mail mode, web mode surveys (1) attain higher response rates among web users, (2) display lower item nonresponse, and (3) elicit more precise values for financial measures. We conclude that, for web-using populations, web mode surveys appear to result in more usable data than mail mode surveys, and these data appear to be of high quality. However, we also find no systematic mode differences in the categorical distributions of responses to items, providing no evidence that pooling data from the two modes is inadvisable.

Introduction

With the proliferation of internet access, fielding surveys over the World Wide Web yields benefits over conventional modes like telephone interviews or paper questionnaires. Web-based surveys can be cheaper and faster to implement than mail questionnaires for large samples. In addition, web-based surveys can be easily customized for subsamples using skip logic, pre-loads, or text "fills," allowing for interaction between respondent and survey in a self-administered mode. While interviewer-administered interviews can also incorporate customization, web interviews are particularly attractive for collecting sensitive economic and financial information, since they preserve respondent privacy. In addition, relative to mail mode, web respondents can more conveniently access online banking records that help them provide accurate responses. Web questionnaires can also use reminder screens to reduce item nonresponse or to offer categorical answer choices when respondents skip free-response financial questions.

However, implementing a survey entirely by web raises concerns about sample selection, since internet access is not randomly distributed. Even among those with internet access, lack of online experience may increase response burden. Furthermore, concerns about internet security may reduce respondents' willingness to participate (Evans and Mathur, 2005), particularly for potentially sensitive financial information. Other features of web mode may influence response characteristics as well.

We investigate how unit response and data quality are affected by the choice of mode or mode-specific design factors. This question is particularly important for studies that may supplement a web survey with an additional mode to address the coverage problem (Rookey et al., 2008). In this paper, we exploit a uniquely designed mode experiment in which respondents with internet access are randomly assigned to take a survey by mail or web.

We find little difference in response rates of completed surveys by mail and web. However, including partial web completions and mode switchers yields a higher response rate for web invitees than for mail invitees. While we find lower item nonresponse on web surveys than mail surveys for near-identical fixed-choice questions, the distributions of the answer choices remain similar across mode. Mode differences in item nonresponse do emerge, however, on financial questions that elicit dollar values, likely due to features of the web questionnaire designed to lower nonresponse and to encourage precise responses. That said, the implied distributions of income responses are similar across mode. This similarity suggests that if researchers prefer precise values to ranges, but in turn prefer ranges to nonresponse, web mode can be leveraged to encourage precise values, and those responses can potentially be used together with mail mode responses in economic analysis.

Background

In this paper, we first examine the impact of mode on unit response and item nonresponse. Next, we turn to mode difference in response types and response distributions. We focus on a household survey of economic behavior containing questions that are typically of interest to economists.

Theories of response—including leverage-saliency theory (Groves, Singer, and Corning, 2000), economic exchange theory (Biner and Kidd, 1994; Porter and Whitcomb, 2003; and Lavrakas, 2008) and social exchange theory (Dillman, 1978; 2007)—suggest that potential respondents are more likely to respond the lower the costs or the personal burden of participating. The relative costs of responding to a web or mail survey are likely to differ depending on the target population. Empirical research on the effect of mode on

response rates provides mixed evidence, with large variation in response rates across studies (see Couper, Blair, and Triplett, 1999, or Fricker and Schonlau, 2002, for example). Studies of mixed-mode surveys have generally found that response rates on mail surveys are higher than those on web surveys (Couper, Blair, and Triplett, 1999; Fan and Shih, 2008). Many of these surveys initially contacted their target web population by e-mail and their target mail population by postal mail, which may contribute to the discrepancy (Kwak and Radler, 2002; Kaplowitz, Hadlock, and Levine, 2004; and Converse, et al., 2008). Indeed, when both groups are initially contacted by postal mail, or when multiple modes of contact are used for web invitees, web response rates are improved relative to mail (Kaplowitz, Hadlock, and Levine, 2004; Messer and Dillman, 2011; Kaplowitz, et al., 2012; and Israel, 2013). In accordance with these findings, our study employs a single mode of contact for all respondents, whether they are assigned mail or web.

Coverage differences between mail and web target samples in studies may also contribute to mode differences in response rates. For populations with little coverage error, studies have found web response rates to exceed mail response rates (Kiernan et al., 2005). For example, in studies of college students, web response rates are greater (Shih and Fan, 2008), and random assignment of survey mode has yielded comparable sample demographics across modes (McCabe et al., 2002; Miller et al., 2002). Our study randomly assigns mode and collects a wealth of information on demographics, personality, and cognition, allowing us to compare our two groups on many dimensions.

This study focuses on older Americans. While data from several years ago have shown that web surveys of older Americans may result in selective samples (Couper, et al., 2007; Fox, 2005; Horrigan, 2008; Hsu, Fisher, and Willis, 2011), nearly 60 percent of adults age 65 and over reported going online in 2013 (Pew Research Center, 2014). As would be expected with improved internet access among the older population, at least one recent study has shown that web mode can be a useful tool for collecting data among this population (de Benardo and Curtis, 2012). As such, increasing numbers of surveys are moving or exploring a move to internet mode. While internet access is not yet universal, it is useful to understand the differences that may be introduced between mail and internet modes of self-administered surveys. This study presents evidence from one survey.

In our study, we combine a mode experiment in which we randomly assign mode to the target population with the option for nonresponders to switch to the other mode. These respondents can be interpreted as switching from a nonpreferred mode to a preferred mode, given the choice. Some papers have theorized that offering respondents their preferred mode will increase response rates (Groves, Singer, and Corning 2000; and Dillman, Smyth, and Christian 2009). Our findings seem to corroborate these theories. Additionally, all initial contact in our study is sent by postal mail, thereby avoiding the potential that survey e-mails are marked as spam or otherwise ignored.

Studies generally find that, conditional on unit response, web mode tends to yield item nonresponse at least as favorable as mail mode. Some studies find little difference in item nonresponse between mail and web surveys, particularly for fixed-choice questions (Denscombe, 2009; and McCabe et al., 2002). Other studies, however, find that web surveys have lower item nonresponse than mail surveys, both in experiments (Messer, et al., 2012) and nonrandom mode assignment (Shin, et al., 2012; and Kwak and Radler, 2002). Like these, our study finds lower item nonresponse in the web mode of the survey.

Other data quality measures include distributions of responses to fixed choice items, including the incidence of extreme answers. In particular, studies have found that telephone respondents are more likely to give extreme answers, particularly on positive responses, than other modes (de Leeuw, 2011; and Ye, Fulton, and Tourangeau, 2011). We focus more broadly on distributions of answers among fixed choices.

A Mode Experiment

Mixed-mode surveys offer two or more modes to respondents (de Leeuw, 2005; and Dillman, Smyth, and Christian, 2009). Most assign mode nonrandomly, for example, by administering web surveys to those who are internet users and mail or phone interviews to those who are not, and therefore lack cross-mode comparison groups. Indeed, disentangling mode differences from coverage differences is difficult without a mode experiment that randomly assigns otherwise comparable individuals to different modes.

The uniquely designed, mixed-mode Cognitive Economics (CogEcon) Study, fielded in 2008, presented an ideal opportunity to conduct one such experiment.[1] By randomly assigning potential respondents who previously reported using the internet regularly to either the mail or web modes of the survey, this study helps shed light on mode differences inherent to web- versus mail-based surveys and survey data. Because the CogEcon mail and web instruments were designed in tandem, balancing alternate-mode comparability with the desire to exploit unique features of web surveys, CogEcon's design resulted in a particularly useful experiment for understanding mode differences in similarly designed studies.

The web survey took advantage of some of the unique features of web surveys, relative to mail surveys: automatic skip logic, reminder screens for skipped questions, and the ability to ask some questions sequentially. As such, our "mode experiment" should be thought of as allowing us to test for mode differences resulting from invitation to, and/or completion of, very similar, but not identical, web and paper surveys. We specify in each analysis section which mode differences we are attempting to examine.

Study Design

The CogEcon study was originally developed to help understand the cognitive bases of economic decisionmaking of older adults. Its sample was drawn from participants from a study of the cognitive abilities of older Americans, the Cognition and Aging in the USA Study (CogUSA). The CogUSA sample, a national sample of older Americans, was recruited using random-digit dialing screener calls from a sample frame that mirrors that of the Health and Retirement Study.[2] Individuals aged 50 or older in 2007 were eligible for the CogUSA study, as well as their spouses and partners. In all, 1222 respondents were invited to participate in the CogEcon Study in 2008.[3]

CogUSA collected socio-demographic, health, cognition, and personality data in a single mode (the CogUSA variables), whereas the economic and financial data were collected for CogEcon either by web or mail, with randomized mode assignment for those with web access. Because each CogUSA variable

---

[1] The Cognitive Economic Study (CogEcon) is sponsored by the National Institute on Aging (grant number NIA P01 AG026571) and is conducted by the University of Michigan.

[2] The CogUSA Study was sponsored by the National Institute on Aging (grant number R 37 AG007137, PI: McArdle). Information on the Health and Retirement Study (HRS) is available at http://hrsonline.isr.umich.edu/.

[3] See the appendix for more information on the CogUSA-CogEcon sample design.

was collected in a single mode, these variables can be used as controls in our analysis without further complicating mode considerations across respondents.

Experimental Design

Web-eligibility was determined using a question from the CogUSA interviews.[4] Respondents who reported that they regularly used the internet were deemed "web-eligible," while all others were considered "web-ineligible." Of the 1222 potential CogEcon respondents, 813—two-thirds—were web-eligible.

The 2008 CogEcon survey included a randomized mode experiment. Of 813 web-eligible CogUSA respondents, 624 were first invited to take a web survey, and 189 were offered mail questionnaires. Web-eligible individuals were randomly assigned to the two modes, while all web-ineligible individuals were assigned to the mail mode, so we will focus our discussion on the two groups of web-eligible respondents.[5]

All potential respondents were sent a pre-notification postcard and an invitation letter with a $25 check that they were free to keep regardless of survey completion, in accordance with survey response theories of economic exchange and social exchange. Respondents assigned to web mode were sent an invitation letter with a URL for an internet interview. Up to three reminder letters were sent to invitees who had not yet completed the web survey. The last of the reminder letters included a copy of the mail questionnaire, offering an explicit opportunity to switch modes. Thirty members of this web-eligible group assigned to web mode eventually submitted a mail questionnaire, either because they requested it from the study, or because they submitted the mail survey sent with the final reminder. Those assigned to mail mode received a copy of the mail survey with the invitation letter and up to three reminders. The invitation letter and reminders each included a new copy of the mail survey.

This paper compares web-eligible respondents assigned to either web or mail mode to investigate mode effects in the experiment embedded in the CogEcon study. We first seek to establish that this mode experiment is valid. In particular, using data gathered during preliminary single-mode interviews, we establish that the experimental design allows us to compare mode effects across comparable samples of individuals. We then turn to investigating unit response rates by mode, using several different definitions of response rates. Next, we analyze data quality with respect to item nonresponse as well as response distributions. We consider fixed-choice (or quasi-fixed choice) questions that are nearly identical across modes, as well financial questions eliciting dollar values that take greater advantage of design innovations possible on a web instrument. [6] Ultimately, we find qualitatively similar results whether or not we control for demographic, cognition, and personality variables collected in preliminary single-mode interviews.

---

[4] Respondents were asked: "Do you regularly use the World Wide Web, or the internet, for sending and receiving e-mail or for any other purpose, such as making purchases, searching for information, or making travel reservations?" Those who answered yes were web-eligible, while all other responses (including refusals) were web-ineligible.
[5] This design also allows for comparisons between both web-eligibility groups but within mode (Hsu et al., 2011).
[6] The notable differences were reminder screens on the web and a single-question-per-page design on the web. These design differences were infeasible in the mail mode but were deemed important to reduce item nonresponse and respondent burden in spite of introducing differences across mode.

Characteristics of Survey Invitees and Validity of the Mode Experiment

As discussed, web-eligible invitees were randomly assigned to release 1 (web) or release 2 (mail) for the survey. Table 1 shows summary statistics for each of the target groups; column (5) displays the difference in means across the groups, and the last column displays the *p*-value from a two-tailed t-test. We included some standard demographic variables (age, sex, coupleness status, education, indicator for fair/poor health), as well as additional psychological variables that we hypothesized may affect response: a standardized Number Series score (a cognition test, adapted from the Woodcock Johnson III (WJ III), that measures quantitative analytical reasoning ability), a standardized Verbal Analogies score (a cognition test from the WJ III that measures the ability to reason using lexical knowledge), an indicator for low mental status (designed to capture cognitive impairment), measures of Big Five personality traits, a measure of Need for Cognition, an indicator for military experience, indicators for being politically conservative or liberal (we omitted the moderate as the comparison group), and an indicator for frequent attendance of religious services. Note that these characteristics were all measured in the same mode for both groups. They were measured over the telephone for most variables, and in face-to-face interviews for the cognition and personality variables.

The two groups look very similar, with a few exceptions. Those assigned to the web group are somewhat more likely to be female, more likely to have a low mental status score, and more likely to have higher scores for the Big Five personality measure for neuroticism. They are also more likely to be politically conservative and to frequently attend religious services. Of these characteristics, only neuroticism is statistically significant at the 5 percent level or better. Statistically, these differences do not mean that assignment was not truly random, but we should be aware that the two target samples do differ in ways that would generally be unobservable to researchers using data from studies without psychological measures. These differences may influence the results of our item response analyses if individuals with some of the characteristics that are more or less common in one mode tend to respond differently to survey items.

Results

In our analysis, we first compare the effect of the invitation treatment on unit response. That is, we examine the impact on various definitions of unit response of inviting individuals to the mail survey versus the web-based survey. We consider different definitions of unit response, including combinations using fully completed surveys in the assigned mode, partially completed web surveys, as well as surveys completed in a different mode than assigned.

Turning to data quality, we then consider responses to two groups of questions, fixed-choice and financial questions. In this section, we compare the impact of the mode of survey completion for those respondents who submitted at least some responses via the mode to which they were originally invited.[7]

First, we analyze the quality of seven fixed-choice questions that were asked of all respondents and are nearly identical in both modes. These questions used identical wording and the layout was as identical as

---

[7] We exclude responses from "mode switchers" in these analyses because they do not clearly fall into either response mode. That is, we compare data from individuals who submitted responses in the mode to which they were originally assigned. "Mode switchers" are not included in either comparison group.

possible, given the mode difference. The differences that we think may be meaningful are that the web survey generally presented only one question per page (versus several for the mail survey), and the web survey displayed "soft-check" reminders after skipped questions. While web respondents could choose to skip a question, the reminder screens increased the burden of doing so relative to the mail mode. We examine item nonresponse counts and response distributions of these questions. Given the existing literature on similar mixed-mode surveys, we expect to see higher rates of item nonresponse on the mail surveys. If web respondents are more likely than mail respondents to use satisficing techniques to avoid the reminder screens, we may also see different response distributions when compared to the mail mode, where respondents may skip questions without being faced with a reminder screen.

Second, we investigate three financial questions that elicit dollar values. The mail instrument offered respondents two ways to answer each financial question side by side: a text box to fill in a dollar value, or a fixed choice of ranges. On the web instrument, the text box was offered first, with the range option presented on the following screen if the respondent skipped the text box. Therefore, the web instrument was designed to nudge respondents toward reporting an exact value instead of a range. We analyze how these differences in the instruments for each mode influenced item nonresponse, the choice of reporting in values versus ranges, as well as the implied distribution of values.

Unit Response Rates

To analyze the effect of mode on response rate, we restrict our analysis to the web-eligible respondents, who were randomly assigned to web or mail survey modes, 813 individuals in total. As discussed above, all respondents were invited using a mailed letter containing either instructions for accessing the questionnaire over the web, or a mail questionnaire. The 813 invitees can be categorized into the following response categories:

1. Completed survey and submitted in the mode to which the respondent was assigned
2. Completed survey and submitted in a different mode than the assigned mode
3. Partially completed web survey
4. No survey submitted or no responses submitted on web

Group 2 respondents, hereafter "mode switchers," were initially invited to take a web survey. However, these respondents either requested and subsequently submitted the mail survey, or submitted the mail survey that was mailed with the final reminder letter.

Response rates calculated using different subsets of these groups are presented in Table 2. If we consider only fully completed surveys (the minimum response rate, AAPOR RR1) in the mode the respondents were assigned to, the response rates are very similar; 78.8 percent for web and 77.2 percent for mail.[8] A chi-squared test reveals no statistically significant difference between the two ($p = 0.640$).

Adding those who partially completed web surveys (that is, respondents who logged into the web survey and completed at least one substantive question but did not click the "submit" button at the end of the

---

[8] Because the target sample was drawn from an existing survey with the same eligibility rules, the denominator of our response rates are the sum of completed interviews (both those completed in the assigned mode and those completed in the other mode), partial interviews, and refusals.

survey) for RR2, the web response rates increases slightly, but a chi-squared test again does not reveal statistically significant differences in the two response rates ($p = 0.114$).

There were 30 mode switchers. Comparing the response rates of mail invitees to web invitees, where these mode switchers are included as unit respondents in the web invitee group, those assigned to the web survey had a response rate of 83.7 percent, and were more likely to respond than mail invitees ($p = 0.044$). Including both partial surveys as well as these mode switchers, the response rate of web invitees increases to 87.2 percent, nearly 10 percentage points higher than mail invitees. This difference is precisely measured ($p = 0.001$).

In Table 1, we showed that web and mail invitees looked similar, with a few exceptions. To determine whether mode effects on response rates still emerge when controlling for sample characteristics, Tables 3 and 4 present results from logistic regression models with a response indicator on the left hand side and assigned mode and demographic, cognitive, personality, and other variables on the right. We use indicator variables for five-year age categories, since effects of age may not be linear. While valid measures for some personality and cognition variables are missing for a small number of respondents, results with fewer covariates are qualitatively similar.

In Table 3, we examine the effect of invitation mode on the most conservative definition of response rate, where response is defined as having submitted a survey in the assigned mode. The outcome variable is equal to 1 if the respondent submitted a fully complete survey in the assigned mode and 0 otherwise. Average marginal effects are reported. Whether we estimate a univariate model (column (1)), include demographic controls (column (2)), or include demographic, cognition, personality, and other variables (column (3)), the estimates never show a statistically significant mode effect.[9] In a parallel analysis in which both partial web surveys and completed surveys are treated as response, there was also no mode effect.

By contrast, a strong mode effect is revealed when adding all marginal submissions, including both partial respondents to the web survey and mode switchers in the respondent group. In Table 4, the outcome variable is equal to 1 if a respondent submitted a complete or partial survey in either mode and 0 otherwise. The average marginal effect for the univariate regression shows that assignment to the web mode treatment increases the probability of response by 0.088; after including all controls, the average marginal effect of web assignment is 0.096. The mode effect is precisely measured for all three specifications ($p < 0.001$).[10]

Next, we consider the characteristics of the resulting samples of web and mail respondents. Columns (1) and (2) of Table 5 display summary statistics for respondents who submitted any survey responses, by mode of invitation. Column (2) displays the characteristics of respondents who were assigned to the web, and includes respondents who completed the web survey as well as both types of marginal respondents:

---

[9] Being in good health or in a couple, however, does increase the likelihood of completing a survey.

[10] For brevity, results for variables other than "web mode" are reported in the Appendix. The only notable findings from these results are that the effect of health is only statistically significant when cognition/personality are included and, in contrast to results from Table 3, the effect of coupleness status is not statistically significant in Table 4.

partial respondents and switchers. Column (3) displays p-values for tests of differences by mode of invitation. These are calculated using results from two-tailed t-tests for continuous variables and chi-squared tests for categorical and indicator variables. The two groups look very similar on most variables. In this table, which compares *respondents* by invited mode, mail respondents have lower values of neuroticism (difference of about 0.2 standard deviations, $p = 0.026$), are 7.4 percentage points more likely to be politically conservative ($p = 0.093$), and are 8.7 percentage points more likely to attend religious services ($p = 0.064$). Results in these columns closely parallel to the results in Table 1, which compares web versus mail *invitees*, and suggest selection patterns into response are similar by mode.

Columns (4) and (5) compare complete submissions in the assigned mode (N = 638) with the marginal submissions from respondents assigned to the web (N = 52). The marginal respondents are generally more likely to be older, not coupled, and less healthy, and also display some differences in cognitive measures, relative to respondents who completed the survey in their assigned mode. The differences between these two groups highlight the importance of including marginal submissions in final datasets whenever possible: the marginal respondents look different on a number of measures. Thus, a survey that attempts to be representative of its invitees ought to include such marginal respondents.

In sum, response rates for fully completed surveys in the assigned mode are very similar for web and mail. When we include partially completed surveys or mode switchers, web invitees respond at a statistically significant higher rate than mail invitees. Similar patterns emerge when we control for demographic, cognitive, and personality variables. Including data from both partially completed surveys and surveys from mode switchers may help reduce mode-related selection bias in mixed-mode studies.

Item response

In this section, we compare the effect of mode of survey completion induced by invitation to that mode on item non-response rates, distribution of responses, and use of range versus precise response options. As such, we report analyses that compare responses from web-eligible mail invitees who submitted mail surveys with web-eligible web invitees who submitted responses via the web survey. Both complete and partial web survey submissions are considered. Mode switchers are excluded from the analyses presented because they do not clearly fall into either response category we are considering.

Near-identical fixed-choice questions

Item nonresponse

Seven items on the survey were asked of all respondents, using identical phrasing, response choices, and layout in both modes. Due to the web-based survey design advantages used in the study (web reminder screens and one question per page layout), we expect that the item nonresponse rates will be lower on the web than for the mail questionnaires. Below, we first establish that item nonresponse rates do differ by mode; we then use a multivariate analysis to examine whether the differences remain after controlling for other covariates. We verify that item nonresponse is not related to covariates other than the random mode assignment, since such correlations could bias inference using these data.

The set of near-identical items across the two modes examined in this section include self-assessments of risk tolerance, financial readiness, and mathematics skills, as well as objective questions, like whether or not the respondent has a will. See the Appendix for a description of the seven items included in the analysis.

9

We created a count of the number of these seven questions that were skipped by each respondent and tested for differences in means, variances, and distributions between the mail and web modes. We compared two different samples of web respondents to the mail respondents: web respondents who had completed and submitted the survey ("Web completes"), and web respondents who had partially completed but not submitted the web survey ("Web including partials"). Web invitees who switched to mail mode are excluded from this analysis.[11]

The first part of Table 6 reports the distribution of items skipped by respondent. The first column is for mail respondents. The second column includes only completed web surveys; the third also includes partially-completed web surveys. Mail respondents were more likely to skip questions, and if they skipped any questions, they tended to skip more questions than web respondents. Including partially completed web surveys in the web sample makes the distribution more similar to the mail distribution, but differences remain. Fisher's exact tests for equality of distributions between mail respondents and each of the web samples are reported below the cell percentages. These tests indicate that the distributions between the mail and web surveys are statistically significantly different ($p < 0.001$ for the sample excluding partial web responses and $p = 0.026$ for the sample including partials). A Brown-Forsythe test for equality of variances also indicates statistically significant differences between the mail sample and the two web samples ($p < 0.001$ for the sample excluding partial web responses and $p = 0.026$ for the sample including partials).[12]

As a robustness check, we estimate multivariate negative binomial regression analyses of the count of items missing on an indicator for web mode and other demographic and psychological covariates. Again, we find that web mode is highly predictive of lower item nonresponse; results are reported and discussed in the Appendix. These results are consistent with our earlier findings that web and mail invitees were similar and provide evidence that mode differences were not driven by some other variable that might be correlated with web mode.

Overall, being invited to complete a mail questionnaire is the most important predictor of item nonresponse in our analysis. As mentioned above, this result is likely due, at least in part, to the reminder screens that are shown if respondents skip items on the web version of the survey. Given that there is no other observable variable that is strongly predictive of item nonresponse, it seems unlikely that the differential item nonresponse patterns on web and mail will introduce bias in analyses using the survey data.

---

[11] Some of these respondents may have partially completed a web survey prior to submitting the mail survey. In this case, only the data from the more complete survey (generally the mail mode for these "switchers") were retained.

[12] Chi-squared tests, while common in the literature, are not appropriate here because of the presence of several cell counts under five. T-tests and ANOVA are likely not appropriate due to the non-normality of the distribution of the number of items skipped, as well as the differences in the distributions. The Brown-Forsythe F-tests are more robust than Bartlett's F-test and Levene's F-test for equality of variances under non-normality, skewness, and kurtosis.

Distributions of response choices

To further examine mode differences in the CogEcon study, we next analyze the distributions of responses to the same seven questions considered above.

Table 7 presents summary statistics of each of the questions, along with ANOVA F-statistics and their associated p-values. Results from Mann-Whitney-Wilcoxon and Pearson's chi-squared tests of equality of distribution, Brown-Forsythe F-tests of equality of variation, Welch's T-test, and Fisher's Exact tests (when appropriate) yield similar results; we will indicate which results differ from the ANOVA.

There are three rows for each question; the first summarizes the responses given by web-eligible mail invitees ("Mail"), the second the web-eligible web invitees who completed a web survey ("Web completes"), and the third web-eligible web invitees who submitted at least some substantive responses via the web survey ("Web incl. partials").

Panel A displays results for the self-assessments of risk tolerance (question 4) and financial planning quality (question 5), which were reported on a 0 to 10 scale. The summary statistics for question 4 look quite similar across groups, with the exception of web groups appearing to have slightly higher means than the mail group. However, the differences are small and not statistically significant (ANOVA $p = 0.44$ and $p = 0.58$ for comparisons of mail with each of the web groups). Overall, there is nothing to indicate that question 4 was answered differently in the web versus mail survey mode. Results in Panel A also show no evidence that question 5 was answered differently in the web versus mail survey mode ($p = 0.78$ and $p = 0.85$).

Panel B displays results for three 6-point Likert items for self-rated understanding the stock market (question 10), self-rated math ability (question 11), and the desirability of safe investments (question 13). Again, the ANOVA yields no statistically significant difference between responses to each question by mode. In question 10, three tests out of twelve revealed a marginally significant difference in distributions of the responses. Here, Pearson's chi-squared test and Fisher's exact test for equality of distributions found marginally different distributions between web and mail.[13] For questions 11 and 13, however, all the other tests yield results like the ANOVA—that is, there are no statistically significant differences between web and mail.

Panel C presents results for questions on whether the respondent has a will (question 107) or life insurance (question 108). Here, again, none of the test statistics provide any evidence that respondents in the web mode answered these questions differently than those in the mail mode.

Overall, there is no evidence for systematic mode differences for responses to the questions that were presented nearly identically across modes. Only one item (question 10) showed any evidence of differences in response distributions, but the findings were of marginal significance and only in three of twelve closely related tests performed on the data from this item. It seems likely that this is simply a type 1 error; however, it may be that inference using cross-mode pooled data from this item should be interpreted with caution.

Dollar value questions with mode innovations
In this section, we investigate whether our survey yielded different response patterns by mode. As detailed below, the study was designed to exploit some of the technological innovations of the web to

---

[13] Comparing "web completes" with mail, $p = 0.1$ and $p = 0.119$ for Pearson's chi-square and Fisher's exact tests, respectively; comparing web (including partial submission) with mail, the $p = 0.073$ and $p = 0.089$.

increase exact values and reduce nonresponse to potentially sensitive financial questions. Therefore, we should interpret any differences by mode as a combination of pure mode effects, effects from questionnaire design innovations on the web survey, as well as the recruitment regimen.

Many of the questions on the survey required a dollar value answer. For such questions that were asked of all respondents, respondents were given the option of either providing a value or choosing one of 10 ranges. Ranges are offered in order to elicit some sort of response from respondents who may be unwilling or unable to provide a precise value, thereby reducing item nonresponse.

On the web instrument, a respondent would see a screen with a text box to enter the dollar value. If he left the answer blank and clicked "next," he would see the option of choosing a range. Therefore, if a respondent entered an exact value when prompted, he would never see the range option. On the paper questionnaire, exact value and range options are all on the same page, side by side. These design differences were allowed to leverage the multiscreen capabilities of web, as precise values are preferred to range responses for economic research, while range responses are preferred to nonresponse.

To assess mode differences for these dollar-value questions, we analyze three questions that were posed to all respondents: total household income, how much was spent on food at home, and how much was spent on food away from home. Table 8 presents the proportions of responses to the three questions that were reported as exact values, ranges, or nonresponses, by mode.

First, item nonresponse is quite low: 5.5 percent for mail responders and 3.1 percent for complete web responders. The difference in nonresponse rates is statistically significant ($p = 0.017$). However, if we include partial responders, the mode difference is no longer statistically significant ($p = 0.602$).

More strikingly, the vast majority of web respondents reported exact values, while most mail respondents utilized ranges. A chi-square test indicates that the two response-type distributions are statistically different ($p < 0.0001$). The patterns are very similar whether or not partial responders are included or if we look at each variable individually.

In Table 9, logistic regression analyses of item nonresponse (column (1) for complete submissions, and column (2) adding partial submissions), and, conditional on response, whether or not the respondent reports an exact value (completes only in column (3), and completes and partials in column (4)), are reported. Even controlling for demographic, personality, cognition, and other variables for each question, and estimating robust standard errors adjusted for respondent-level clustering, we see a significant negative relationship between web mode and item nonresponse (that is, lower item nonresponse on web) when considering complete submissions only. Likewise, web mode is strongly associated with providing exact value responses, increasing the probability by over 30 percentage points, precisely measured, over mail mode.

In Table 10, we convert all dollar value responses to the income question to their appropriate ranges and see if the distribution of ranges differs across mode. The two distributions are statistically indistinguishable under either a Pearson chi-square test or a Fisher's exact test ($p = 0.164$ and $p = 0.194$, respectively). For the two food questions, over 97 percent of all respondents select the lowest range for food at home and 93 percent for food away from home; Fisher exact tests for these two variables also cannot distinguish mode differences in the distribution ($p = 0.741$ and $p = 0.487$, respectively).

13

In sum, item nonresponse rates for these three dollar-value questions are comparable across mode. Web respondents tend to report exact values, whereas mail respondents tend to report ranges. However, both modes yield similar proportions of responses in each range, providing no evidence that contraindicates conducting analyses using pooled data from both modes.

Discussion

The Cognitive Economics study provides a unique opportunity to understand mode effects in a survey administered using web and mail questionnaires. In contrast to many other studies of mode effects, our study incorporates randomized mode assignment among individuals who report web access, pre-notifications and invitation by mail for both modes, and a wide variety of question types, including financial questions requiring a dollar-value response. The study focuses on older Americans, a group that has historically had relatively lower levels of web usage relative to college students and other younger populations typically used for these studies.

We find little difference in response rates of completed surveys by web and mail. However, including partial web completions and mode switchers yields a response rate higher for web invitees than for mail invitees.

Item nonresponse on near-identical fixed-choice questions was lower for the web mode of the survey than the mail mode. Additionally, the distributions of the answer choices in the set of near-identical fixed-choice questions did not appear to differ systematically by mode.

Item nonresponse on financial questions was similar by mode. Response type—exact value versus range—differed dramatically by mode. However, this was more likely due to the design of the web questionnaire, which only offered respondents the option to use a range if they declined to provide an exact value. Pooling both response types together, the implied distributions of responses on income look very similar between mail and web respondents. Therefore, if researchers prefer precise values to ranges, but are willing to use ranges when the alternative is nonresponse, our results provide evidence that the two response types exhibit similar patterns and could potentially be used together in economic analysis.

References

Cognition and Aging in the USA Study. 2011. Waves 1, 2, and 3 datasets. Produced by the University of Michigan for the University of Southern California with funding from the National Institute on Aging (grant number NIA R 37 AG007137, PI: McArdle).

Cognitive Economics Study. 2012. CogEcon Public Core Release (Version 1.0) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA P01 AG026571), Ann Arbor, MI.

Couper, Mick P., Johnny Blair, and Timothy Triplett. 1999. "A Comparison of Mail and E-mail for a Survey of Employees in U.S. Statistical Agencies." *Journal of Official Statistics* 15(1): 39–56.

Couper, Mick P., Arie Kapteyn, Matthias Schonlau, and Joachim Winter. 2007. "Noncoverage and nonresponse in an Internet survey." *Social Science Research* 36: 131-148.

de Bernardo, Dana Huyser and Anna Curtis. 2012. "Using Online and Paper Surveys: The Effectiveness of Mixed-Mode Methodology for Populations Over 50" *Research on Aging* 35(2): 220-240.

de Leeuw, Edith. 2005 "To Mix or Not to Mix Data Collection Modes in Surveys." *Journal of Official Statistics* 21:233–55

de Leeuw, Edith, Joop Hox and Annette Scherpenzeel. 2011. "Mode Effect or Question Wording? Measurement Error in Mixed Mode Surveys." *Proceedings of the Survey Research Methods Section, American Statistical Association*, *2011,* 5959-5967.

Denscombe, Martyn. 2009. "Item non-response rates: a comparison of online and paper questionnaires." *International Journal of Social Research Methodology* 12(4): 281-291.

Dillman, Don A., Jolene D. Smyth, and Leah Melani Christian. 2009. *Internet, Mail, and Mixed Mode Surveys: The Tailored Design Method 3rd ed.* Hoboken, NJ: John Wiley & Sons

Evans, Joel R., and Anil Mathur. 2005. "The value of online surveys." *Internet Research* 15: 195-219.

Fricker, Ronald D., and Matthias Schonlau, Jr. 2002. "Advantages and disadvantages of Internet research surveys: Evidence from the literature." *Field Methods* 14(4): 347–67.

Frippiat, Didier, Nicolas Marquis, and Elizabeth Wiles-Portier. 2010. "Web Surveys in the Social Sciences: An Overview." *Population, English edition* 65: 285-311.

Groves, Robert M., Eleanor Singer, and Amy Corning. 2000. "Leverage-Saliency Theory of Survey Participation: Description and an Illustration." *Public Opin Q (2000)* 64(3): 299-308 doi:10.1086/317990

Horrigan, John B. 2008. "Home Broadband Adoption 2008." Retrieved September 22, 2014 (http://www.pewinternet.org/Reports/2008/Home-Broadband-2008.aspx).

Hsu, Joanne W., Gwenith G. Fisher, and Robert J. Willis. 2011. "Internet Access and Cognitive Ability: An Analysis of the Selectivity of Internet Interviews in the Cognitive Economics Survey." *Essays on Aging and Human Capital,* Joanne W. Hsu, Ph.D. dissertation, University of Michigan.

Israel, Glenn D. 2013. "Combining Mail and E-Mail Contacts to Facilitate Participation in Mixed-Mode Surveys." *Social Science Computer Review* 31(3): 346-358.

Kaplowitz, Michael D., Timothy D. Hadlock, and Ralph Levine. 2004. "A Comparison of Web and Mail Survey Response Rates." *Public Opin Q (2004)* 68(1): 94-101 doi:10.1093/poq/nfh006

Kaplowitz, Michael D., Frank Lupi, Mick P. Couper and Laurie Thorp. 2012."The Effect of Invitation Design on Web Survey Response Rates." *Social Science Computer Review* 20(3): 339-349.

Kiernan, Nancy Ellen, Michaela Kiernan, Mary Ann Oyler, and Carolyn Gilles. 2005. "Is a Web Survey as Effective as a Mail Survey? A Field Experiment Among Computer Users." *American Journal of Evaluation* 26(2): 245-252, doi:10.1177/1098214005275826

Kwak, Nojin and Barry Radler. 2002. "A Comparison Between Mail and Web Surveys: Response Pattern, Respondent Profile, and Data Quality." *Journal of Official Statistics,* 18(2): 257-273.

McCabe, Sean Esteban, Carol J. Boyd, Mick P. Couper, Scott Crawford, and Hannah D'Arcy. 2002. "Mode effects for collecting alcohol and other drug use data: Web and U.S. mail." *Journal of Studies on Alcohol* 63(6): 755–761.

Messer, Benjamin L. and Don A. Dillman. 2011. "Surveying the General Public over the Internet Using Address-Based Sampling and Mail Contact Procedures." *Public Opin Q (Fall 2011)* 75(3): 429-457 doi:10.1093/poq/nfr021

Messer, Benjamin, Michelle Edwards, and Don Dillman. 2012. "Determinants of Item Nonresponse to Web and Mail Respondents in Three Address-Based Mixed-Mode Surveys of the General Public." *Survey Practice* 5(2).

Miller, Elizabeth T., Dan J. Neal, Lisa J. Robers, John S. Baer, Sally O. Cressler, Jane Metrik, and Alan G. Marlatt. 2002. "Test–retest reliability of alcohol measures: Is there a difference between internet-based assessment and traditional methods?" *Psychology of Addictive Behaviors* 16(1): 56–63.

Pew Research Center. 2014. "Older Adults and Technology Use." Retrieved September 22, 2014 (http://www.pewinternet.org/2014/04/03/older-adults-and-technology-use/)

Rookey, Bryan D., Steve Hanway, and Don A. Dillman. 2008. "Does a Probability-Based Household Panel Benefit from Assignment to Postal Response as an Alternative to Internet-Only?" *Public Opinion Quarterly* 72: 962-984.

Schonlau, Matthias, Arthur van Soest, Arie Kapteyn, and Mick Couper. 2009. "Selection Bias in Web Surveys and the Use of Propensity Scores." *Sociological Methods & Research* 37: 291-318.

Shih, Tse-Hua and Xitao Fan. 2008. "Comparing Response Rates from Web and Mail Surveys: A Meta-Analysis." Field Methods, 20(3): 249–271. doi:10.1177/1525822X08317085

Shin, Eunjung, Timothy P. Johnson, and Kumar Rao. 2012. "Survey Mode Effects on Data Quality: Comparison of Web and Mail Modes in a U.S. National Panel Survey." *Social Science Computer Review*, 30(2): 212-228. doi: 10.1177/0894439311404508

Valliant, Richard and Jill A. Dever. 2011. "Estimating Propensity Adjustments for Volunteer Web Surveys." *Sociological Methods & Research* 40: 105-137.

Ye, Cong, Jenna Fulton, and Roger Tourangeau. 2011. "More Positive or More Extreme? A Meta-Analysis of Mode Differences in Response Choice." *Public Opinion Quarterly* 75(2): 349-365.

Tables

Table 1. Summary statistics of respondents assigned to web and mail

| Variable | Assigned Mail | | Assigned Web | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
	Mean	Obs.	Mean	Obs.	Difference	p-value
Age	60.73	189	61.08	624	-0.35	0.64
Female	0.52	189	0.59	624	-0.07	0.10
Coupled	0.77	189	0.75	624	0.02	0.56
Years of education	14.71	189	14.81	624	-0.10	0.56
Fair/poor health	0.12	189	0.13	624	-0.01	0.69
Need for Cognition: Enjoyment	0.24	188	0.15	622	0.09	0.26
WJ III: Number Series	0.36	188	0.27	623	0.09	0.19
WJ III: Verbal Analogies	0.29	188	0.22	620	0.06	0.43
Low mental status	0.09	189	0.14	624	-0.05	0.09
Big 5 Personality: Extroversion	0.13	188	0.04	621	0.08	0.34
Big 5 Personality: Agreeableness	0.00	188	-0.06	621	0.06	0.50
Big 5 Personality: Conscientiousness	-0.02	188	-0.02	621	0.00	0.98
Big 5 Personality: Neuroticism	-0.15	188	0.02	621	-0.17	0.04
Big 5 Personality: Openness	0.16	188	0.16	621	0.00	0.98
Military	0.15	189	0.17	624	-0.02	0.53
Liberal	0.21	186	0.26	619	-0.05	0.19
Conservative	0.39	186	0.31	619	0.07	0.06
Religious	0.58	188	0.50	622	0.08	0.06

Note: Sample size vary slightly because, in a small number of interviews, some measures were dropped due to either interviewer error or interviewer notes that indicated that measure was invalid. *p*-values are from 2-tailed t-test for difference in means. The data sources are the Cognitive Economics Study and the Cognition and Aging in the USA Study.

Table 2. Response Rates: Completed Surveys and Marginal Submissions

Respondent status includes:		Assigned Mail	Assigned Web	P-value
Submitted surveys	Non-respondents	0.228	0.212	
	Respondents	0.772	0.788	0.640
+ Partially complete web surveys	Non-respondents	0.228	0.176	
	Respondents	0.772	0.824	0.114
+ Submissions by mode switchers	Non-respondents	0.228	0.163	
	Respondents	0.772	0.837	0.044*
+ Partial complete & mode switchers	Non-respondents	0.228	0.128	
	Respondents	0.772	0.872	0.001***
Observations		189	624	

Notes: p-values from chi-squared tests of independence (1 degree of freedom) between respondent/non-respondent status and assigned mode, using each of four different concepts of respondent status. Total number of observations is 813 for each test: 189 assigned mail and 624 assigned Web. The data source is the Cognitive Economics Study. * for $p<0.05$, ** for p<0.01, *** for $p<0.001$

	(1)	(2)	(3)		(1 cont.)	(2 cont)	(3 cont)
Web Mode	0.016	0.019	0.028	WJ III: Number Series			0.01
	(0.034)	(0.033)	(0.034)				(0.023)
Age 50–54		0.019	0.011	WJ III: Verbal Analogies			-0.023
		(0.077)	(0.078)				(0.017)
Age 55–59		0.084	0.083	Low mental status			0.008
		(0.075)	(0.075)				(0.046)
Age 60–64		0.033	0.028	Big 5: Extroversion			-0.025
		(0.078)	(0.079)				(0.015)
Age 65–69		0.078	0.085	Big 5: Agreeableness			0.007
		(0.081)	(0.081)				(0.016)
Age 70–74		0.003	-0.004	Big 5: Conscientiousness			-0.004
		(0.089)	(0.090)				(0.016)
Age 75+		0.059	0.078	Big 5: Neuroticism			0
		(0.082)	(0.083)				(0.017)
Female		-0.018	-0.015	Big 5: Openness			-0.014
		(0.030)	(0.037)				(0.020)
Coupled		0.085**	0.095**	Military			-0.019
		(0.033)	(0.033)				(0.048)
Some College		0.017	0.001	Liberal			0.04
		(0.040)	(0.042)				(0.038)
College Grad		0.042	0.004	Conservative			0.014
		(0.039)	(0.043)				(0.034)
Fair/Poor Health		-0.114**	-0.126**	Religious			-0.005
		(0.038)	(0.039)				(0.030)
Need for Cognition: Enjoyment			0.055**	Pseudo R-Squared	0.000	0.028	0.047
			(0.019)	P-value	0.641	0.021	0.037
				Observations	813	813	793

Note: Results from logit regressions with dependent variable that is one if the respondent submitted a complete survey in the assigned mode and zero otherwise. The data sources are the Cognitive Economics Study and the Cognition and Aging in the USA Study. * for $p<.05$, ** for $p<.01$, and *** for $p<.001$

Table 4. Logit regression of response status, including complete, partial and switched mode submissions (average marginal effects)

	(1)	(2)	(3)
Web Mode	0.088***	0.088***	0.096***
	(0.027)	(0.026)	(0.027)
Demographic controls?	No	Yes	Yes
Personality, cognition and personal characteristic controls?	No	No	Yes
Pseudo R-Squared	0.015	0.039	0.061
P-value	0.001	0.007	0.025
Observations	813	813	793

Note: Results from logit regressions with dependent variable that is one if the respondent submitted a complete survey in the assigned mode, a partially complete web survey, or switched modes and submitted a survey, and zero otherwise. Average marginal effects reported. Covariates added in each column are identical to those added in the corresponding columns of Table 3. Due to space considerations, average marginal effects for these variables are reported in Appendix. The data sources are the Cognitive Economics Study and the Cognition and Aging in the USA Study. * for $p<.05$, ** for $p<.01$, and *** for $p<.001$

Table 5. Characteristics of responders

| | Responders by Assigned Mode | | | Response Type | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
	Mail	Web	p-value	Complete	Marginal submission	p-value
Age	61.13	61.33	0.817	60.95	65.46	0.001
Female	0.51	0.58	0.118	0.56	0.62	0.435
Coupled	0.77	0.76	0.878	0.78	0.56	0.000
Years of education	14.71	14.86	0.448	14.86	14.44	0.165
Fair/poor health	0.10	0.13	0.430	0.11	0.27	0.001
Need for Cognition: Enjoyment	0.23	0.19	0.679	0.23	-0.15	0.004
WJ III: Number Series	0.32	0.29	0.723	0.32	0.05	0.031
WJ III: Verbal Analogies	0.17	0.24	0.484	0.23	0.08	0.273
Low mental status	0.11	0.12	0.787	0.12	0.10	0.643
Big 5: Extroversion	0.12	0.02	0.305	0.05	-0.04	0.567
Big 5: Agreeableness	0.04	-0.07	0.263	-0.05	-0.01	0.782
Big 5: Conscientiousness	-0.03	-0.02	0.975	-0.02	-0.13	0.436
Big 5: Neuroticism	-0.19	0.02	0.026	-0.03	0.06	0.536
Big 5: Openness	0.09	0.17	0.371	0.17	-0.10	0.051
Military	0.16	0.18	0.732	0.17	0.23	0.261
Liberal	0.21	0.26	0.194	0.25	0.21	0.505
Conservative	0.39	0.31	0.093	0.33	0.33	0.956
Religious	0.58	0.49	0.064	0.51	0.48	0.651

Notes: p-values from 2-tailed t-tests for all variables, except: female, coupled, fair/poor health, low mental status, military, liberal, conservative, and religious, for which *p*-values are from chi-squared tests. Marginal submissions include both partial web surveys and mode switchers (all of these were originally assigned to web, but eventually submitted mail survey). Number of observations by column are: (1) 146, (2) 544, (4) 638, (5) 52. The data sources are the Cognitive Economics Study and the Cognition and Aging in the USA Study.

Table 6. Item nonresponse on near-identical fixed choice questions by mode

Number Skipped	Mail	Web completes	Web including partials
0	90.4%	98.8%	95.9%
1	3.4%	0.8%	0.8%
2	4.1%	0.4%	2.7%
3	0.0%	0.0%	0.0%
4	0.7%	0.0%	0.2%
5	0.0%	0.0%	0.0%
6	0.0%	0.0%	0.0%
7	1.4%	0.0%	0.4%
Mean number skipped	0.24	0.02	0.10
Observations	146	492	514
Fisher's Exact P-value:			
H_0:mail and web same distributions		0.000	0.026
Brown-Forsythe test F:		24.412	4.998
H_0: mail and web equal variances		(d.f. 1, 636)	(d.f. 1, 658)
p-value		0.000	0.026

Note: The data source is the Cognitive Economics Study.

Table 7. Item response statistics and distribution tests for fixed-choice questions

	Mean	Median	St. Dev.	Obs.	ANOVA		p-value
Panel A: 11-point self-assessments (0-10)					F-stat	d.f. (1,_)	
Q4: Self-rated risk tolerance							
Mail	5.36	5	2.08	144			
Web completes	5.51	6	2.06	492	0.60	634	0.44
Web incl. partials	5.47	5	2.06	512	0.31	654	0.58
Q5: Self-rated fin. Arrangements							
Mail	3.73	3	2.84	142			
Web completes	3.80	3	2.67	490	0.08	630	0.78
Web incl. partials	3.78	3	2.66	510	0.04	650	0.85
Panel B: 6-point Likert-type items (1-6)							
Q10: Understand stock market							
Mail	3.52	3	1.43	140			
Web completes	3.62	3	1.44	492	0.54	630	0.46
Web incl. partials	3.65	3	1.43	512	0.84	650	0.36
Q11: Pretty good at math							
Mail	2.60	2	1.32	142			
Web completes	2.66	2	1.29	492	0.27	632	0.60
Web incl. partials	2.69	2	1.31	511	0.58	651	0.45
Q13: Only safe investments							
Mail	4.30	5	1.24	141			
Web completes	4.36	5	1.20	491	0.20	630	0.66
Web incl. partials	4.34	5	1.22	510	0.07	649	0.79
Panel C: Objective questions (0-1)							
Q107: Have a will							
Mail	0.67	1	0.47	139			
Web completes	0.72	1	0.45	491	1.19	628	0.27
Web incl. partials	0.72	1	0.45	498	1.20	635	0.27
Q108: Have life insurance							
Mail	0.82	1	0.39	139			
Web completes	0.81	1	0.39	488	0.03	625	0.86
Web incl. partials	0.81	1	0.39	495	0.05	632	0.83

Note: The data source is the Cognitive Economics Study.

Table 8. Type of response to dollar-value questions, by mode, questions 42, 52, and 53

Response type	Mail	Web completes	Web completes + partials
Value	10.96	91.67	89.75
Range	83.56	5.28	5.38
Neither	5.48	3.05	4.86
Total	100	100	100

Note: The data source is the Cognitive Economics Study.

Table 9. Logit regressions of item nonresponse and exact value response on dollar-value questions

	Item non-response		Response with exact value	
	(1)	(2)	(3)	(4)
Sample used	Completes	Completes + Partials	Completes	Completes + Partials
Web mode	-0.027*	-0.006	0.322***	0.323***
	(0.012)	(0.016)	(0.017)	(0.017)
Q52 indicator	-0.053***	-0.051***	0.060***	0.057***
	(0.009)	(0.009)	(0.013)	(0.013)
Q53 indicator	-0.048***	-0.047***	0.064***	0.060***
	(0.010)	(0.010)	(0.014)	(0.014)
Pseudo R2	0.147	0.129	0.574	0.57
Observations	1869	1932	1800	1833

Note: Results from logit regressions with dependent variable that is one if respondent did not answer a question and zero otherwise (columns 1 and 2). For columns 3 and 4, dependent variable is one if respondent answered the question as an exact value. Average marginal effects reported. Robust standard errors clustered at the respondent level. Demographic, personality and cognitive variables included as controls. Due to space considerations, average marginal effects for these variables are reported in the Appendix. The data sources are the Cognitive Economics Study and the Cognition and Aging in the USA Study. * for $p<.05$, ** for $p<.01$, and *** for $p<.001$

Table 10. Distribution of income responses, by mode, question 42

Income (Q42)			Mail	Completed Web	Total
$1	-	$2,500	2	15	17
$2,501	-	$5,000	0	1	1
$5,001	-	$10,000	0	5	5
$10,001	-	$25,000	10	27	37
$25,001	-	$50,000	18	86	104
$50,001	-	$10,000	45	189	234
$100,001	-	$200,000	52	126	178
$250,001	-	$500,000	6	14	20
$500,001	-	$1,000,000	1	5	6
More than $1,000,000			0	4	4
Total			134	472	606

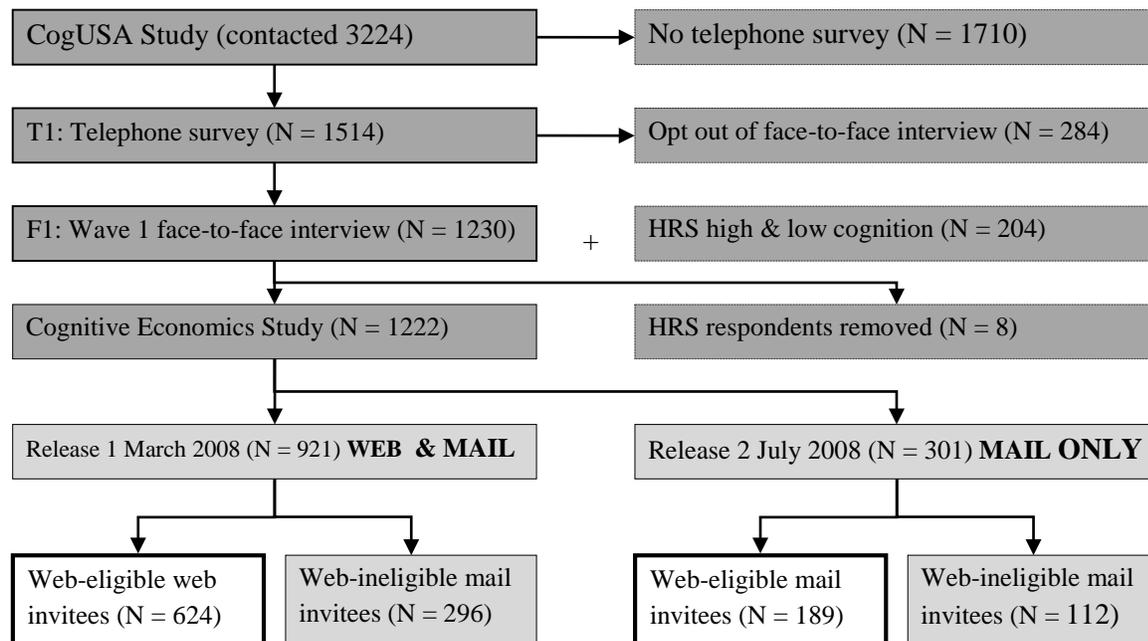*Note:* The data source is the Cognitive Economics Study.

Appendix

Target Sample: CogUSA respondents

The Cognitive Economics study invited participants from a study of the cognitive abilities of older Americans, CogUSA, to complete self-administered mail or web surveys. The CogUSA sample, a national sample of older Americans, was recruited using random-digit dialing screener calls from a sample frame that mirrors that of the Health and Retirement Study.[14] Individuals aged 50 or older in 2007 were eligible for the CogUSA study, as well as all spouses and partners of these "age-eligible" individuals. A total of 1230 respondents completed both the telephone interview and a three-hour face-to-face interview. The CogUSA interviews gathered detailed cognition and personality data and also asked health, socio-demographic, and other background questions from the Health and Retirement Study core interviews.

In total, 1222 respondents were invited to participate in the CogEcon Study in 2008.[15] To avoid conflict with a second wave of the CogUSA data collection, the 2008 CogEcon survey was fielded in two releases. Assignment to the releases was random. For the first release in February 2008, those who had previously indicated that they used the internet regularly were offered the web survey, and those who did not were offered a mail questionnaire. All invitees in the second release, fielded in July, were offered mail questionnaires, regardless of internet use.

Figure 1. Sample Flow Chart: CogUSA and CogEcon Studies

---

[14] Information on the Health and Retirement Study is available at http://hrsonline.isr.umich.edu/.

[15] Eight CogUSA face-to-face interview respondents were deemed ineligible for CogEcon due to their participation in the Health and Retirement Study.

Question wording of near-identical fixed choice questions

- "How would you describe yourself: Are you generally willing to take risks or do you try to avoid taking risks?" (Question 4, 0=Not at all willing to 10=Very willing)
- "If you were to consult a trustworthy and knowledgeable professional financial advisor, how much of your current financial arrangements do you think that the financial advisor would recommend changing?" (Question 5, 0=No change to 10=A total overhaul)
- "I understand the stock market reasonably well." (Question 10, 6-point Likert scale ranging from strongly agree to strongly disagree)
- "I am pretty good at math." (Question 11, 6-point Likert scale ranging from strongly agree to strongly disagree)
- "You should put all of your money into the safest investment you can find and accept whatever return it pays." (Question 13, 6-point Likert scale ranging from strongly agree to strongly disagree)
- "Do you have a written will?" (Question 107, yes or no select one, coded as 1/0)
- "Do you and/or your immediate family have life insurance?" (Question 108, yes or no select one, coded as 1/0)

Discussion of logit regressions of response, split by mode

The first column of Appendix Table 1 displays results from a regression of all invitees where the outcome is 1 for fully complete surveys in the assigned mode; columns (2) and (3) are for those assigned web and mail modes, respectively. Columns (4)–(6) repeat for the outcome that is 1 for all complete and partial submissions.

For completed surveys only, being in a couple and having a high measure for "need for cognition" increases the likelihood of response for web invitees, but there is no precisely measured effect for mail invitees. Having a high verbal cognition score *decreases* the likelihood of response for mail invitees, but for web invitees the effect is not statistically significant. Having low mental status decreases the probability for web invitees but not for mail invitees. For all other variables, the effects are either similar for both modes (poor health), or indistinguishable from zero for both modes (all others).

Now turning to columns (4)–(6) that include partial responders and mode switchers along with completed submissions, the effects for all variables look very similar, with a few exceptions. Low mental status again has a positive effect on mail submissions but a weakly measured negative effect for web submission. Good health and low verbal analogies scores predict mail submissions; both have similar effects for web invitees, though smaller in magnitude and less precisely measured. In contrast to the analysis using completed surveys alone, need for cognition has a similar effect for both modes, as do the remaining variables. For all six specifications, restricting the analysis to variables available for all 813 invitees yields similar results for coupleness, mental status, and health (cognition and personality variables are omitted here due to missing observations in the sample).

Discussion of regression analysis of item nonresponse

Appendix Table 2 presents results from a negative binomial regression of item nonresponse. Negative binomial regression model was chosen over a Poisson model because the variance of the number of missing was much larger than the mean, indicating that a Poisson specification would be inappropriate due to over-dispersion. We conducted regressions with several different subsets of explanatory variables

including web mode indicator only, web mode plus demographics, plus cognitive measures, plus "other." We conducted two separate sets of regressions: one including all completed surveys but excluding mode switchers (column (1)), and one including both partial and complete surveys and excluding only mode switchers (column (2)).

Appendix Table 2 presents only the results from the regressions including the full set of covariates because the different versions yielded largely similar results. In all specifications, web mode was highly predictive of lower item nonresponse (p-values 0.01 or better). For ease of interpretation, we present exponentiated coefficients (and their standard errors) in our results table, or the incidence rate ratios. Holding all else constant, in column (1) the incidence rate ratio of web mode is 0.08. This means that web respondents skipped questions at 8 percent of the rate of mail respondents. In column (2), the incidence rate ratio of web mode is 0.321, meaning that web respondents skipped questions at less than 32.1 percent of the rate of mail respondents.

The coefficient on the "college grad" indicator variable is large and marginally statistically significant (10 percent confidence level), but only in the regression excluding partial web submissions (column (1)). The coefficient on the "religious" indicator variable is also marginally statistically significant, but only in the regression including partial web submissions (column (2)). Here, the incidence rate ratio for religious respondents is 0.464 of the rate of non-religious respondents.

We also conducted negative binomial regressions in which we interacted the web mode indicator with subsets of the other covariates, for both the sample of complete web surveys and the sample including partial web surveys. In these, we used a continuous variable for age to reduce the number of variables and avoid nonconvergence. These regressions resulted in similar, though not identical, findings compared to those only including the main effects: web mode was generally statistically significant at the 10 percent level or better,[16] and low mental status was marginally statistically significant ($p = 0.094$). An interesting exception is that in the regressions containing only web mode and the demographic variables (again, using age as a continuous variable), the incident rate ratio on web mode was similar in magnitude to the regressions presented in Appendix Table 2 but imprecisely estimated. Regression results are available upon request.

---

[16] In specifications including only complete web surveys, web mode was statistically significant at the 1 percent level in all but the regression with only demographics and web mode; in specifications including partial web surveys, the statistical significance was at the 10 percent level for the regressions including cognitive variables, and the 5 percent level for the regressions including the "other" variables.

Appendix Tables

Appendix Table 1. Logit regressions of response, split by mode

	Outcome: complete submission			Outcome: 1 if complete or marginal		
	All invited	Assigned mail	Assigned web	All invited	Assigned mail	Assigned web
	(1)	(2)	(3)	(4)	(5)	(6)
Age 50–54	0.011	-0.084	0.078	-0.015	-0.084	0.031
	(0.078)	(0.118)	(0.093)	(0.068)	(0.118)	(0.085)
Age 55–59	0.083	-0.156	0.186*	0.041	-0.156	0.124
	(0.075)	(0.123)	(0.090)	(0.066)	(0.123)	(0.081)
Age 60–64	0.027	-0.301*	0.149	-0.026	-0.301*	0.077
	(0.079)	(0.135)	(0.093)	(0.070)	(0.135)	(0.084)
Age 65–69	0.083	-0.004	0.132	0.071	-0.004	0.121
	(0.081)	(0.123)	(0.098)	(0.069)	(0.123)	(0.085)
Age 70–74	-0.005	-0.013	0.029	0.029	-0.013	0.071
	(0.090)	(0.130)	(0.108)	(0.076)	(0.130)	(0.093)
Age 75 +	0.078	-0.067	0.134	0.098	-0.067	0.151
	(0.083)	(0.157)	(0.099)	(0.068)	(0.157)	(0.083)
Female	-0.013	-0.041	-0.005	0.003	-0.041	0.012
	(0.037)	(0.068)	(0.043)	(0.032)	(0.068)	(0.035)
Coupled	0.094**	-0.028	0.116**	0.041	-0.028	0.050
	(0.033)	(0.078)	(0.036)	(0.030)	(0.078)	(0.031)
Some college	0.001	-0.018	0.034	-0.026	-0.018	-0.006
	(0.042)	(0.090)	(0.046)	(0.037)	(0.090)	(0.039)
College grad	0.006	0.067	0.008	0.007	0.067	0.002
	(0.043)	(0.103)	(0.046)	(0.039)	(0.103)	(0.040)
Fair/poor health	-0.126**	-0.196*	-0.110*	-0.073*	-0.196*	-0.043
	(0.039)	(0.089)	(0.043)	(0.035)	(0.089)	(0.037)
Need for cognition-enjoyment	0.054**	0.065	0.055*	0.037*	0.065	0.034
	(0.019)	(0.044)	(0.021)	(0.017)	(0.044)	(0.018)
Number Series	0.009	-0.001	0.003	0.001	-0.001	-0.001
	(0.023)	(0.048)	(0.026)	(0.020)	(0.048)	(0.021)
Verbal Analogies	-0.023	-0.068*	-0.006	-0.018	-0.068*	-0.002
	(0.017)	(0.033)	(0.020)	(0.015)	(0.033)	(0.016)
Low mental status	0.009	0.342*	-0.041	-0.032	0.342*	-0.075*
	(0.046)	(0.168)	(0.048)	(0.039)	(0.168)	(0.037)
Personality: Extroversion	-0.025	-0.015	-0.028	-0.020	-0.015	-0.021
	(0.015)	(0.033)	(0.017)	(0.014)	(0.033)	(0.014)
Personality: Agreeableness	0.007	0.046	-0.002	0.009	0.046	-0.001
	(0.016)	(0.034)	(0.019)	(0.014)	(0.034)	(0.016)
Personality: Conscientiousness	-0.003	0.003	0.003	-0.009	0.003	-0.007
	(0.016)	(0.039)	(0.018)	(0.014)	(0.039)	(0.015)

Appendix Table 1, continued. Logit regressions of response, split by mode

	Outcome: complete submission			Outcome: 1 if complete or marginal		
	All invited	Assigned mail	Assigned web	All invited	Assigned mail	Assigned web
	(1)	(2)	(3)	(4)	(5)	(6)
Personality: Neuroticism	0.001	-0.015	0.005	0.001	-0.015	0.003
	(0.017)	(0.037)	(0.019)	(0.015)	(0.037)	(0.016)
Personality: Openness	-0.014	-0.075	0.006	-0.019	-0.075	-0.004
	(0.020)	(0.042)	(0.022)	(0.017)	(0.042)	(0.018)
Military	-0.017	0.074	-0.041	0.016	0.074	-0.009
	(0.048)	(0.103)	(0.054)	(0.043)	(0.103)	(0.046)
Liberal	0.040	0.075	0.031	0.028	0.075	0.020
	(0.038)	(0.085)	(0.043)	(0.034)	(0.085)	(0.035)
Conservative	0.013	-0.006	0.015	0.011	-0.006	0.020
	(0.034)	(0.071)	(0.039)	(0.030)	(0.071)	(0.032)
Religious	-0.006	-0.003	0.005	-0.027	-0.003	-0.014
	(0.030)	(0.063)	(0.034)	(0.026)	(0.063)	(0.028)
Pseudo R2	0.046	0.163	0.074	0.043	0.163	0.060
Observations	793	183	610	793	183	610

Note: Results from logit regressions with dependent variable are 1 if the respondent submitted a complete survey in the assigned mode, and 0 otherwise (columns (1)–(3)). In columns (4)–(6), dependent variable is 1 if respondent submitted a complete survey, a partially complete web survey, or switched modes and submitted a survey, and 0 otherwise. The data sources are the Cognitive Economics Study and the Cognition and Aging in the USA Study. * for $p<.05$, ** for $p<.01$, and *** for $p<.001$

Appendix Table 2. Negative binomial regressions of number of items missing on web mode indicator and covariates

	(1) Completes only	(2) Including partials		(1 cont.) Completes only	(2 cont.) Including partials
Web mode	0.090***	0.405*	Low mental status	3.198	1.627
	(0.048)	(0.149)		-2.266	-0.853
Age 50-54	0.807	1.896	Big 5 Personality:		
	(0.937)	(2.036)	Extroversion	0.73	1.043
Age 55-59	1.33	1.44		(0.214)	(0.201)
	(1.519)	(1.556)	Agreeableness	0.839	1.027
Age 60-64	1.027	1.132		(0.211)	(0.205)
	(1.324)	(1.296)	Conscientiousness	0.893	1.227
Age 65-69	2.664	2.861		(0.242)	(0.248)
	(3.253)	(3.179)	Neuroticism	1.101	1.31
Age 70-74	2.96	2.838		(0.312)	(0.265)
	(3.856)	(3.358)	Openness	0.668	0.778
Age 75 +	0.735	1.056		(0.230)	(0.183)
	(1.122)	(1.347)	Military	0.739	0.523
Female	1.5	1.009		(0.625)	(0.353)
	(0.899)	(0.432)	Liberal	0.889	0.741
Coupled	1.483	0.729		(0.575)	(0.339)
	(0.959)	(0.300)	Conservative	0.91	0.853
Some college	8.493	1.583		(0.529)	(0.360)
	(9.767)	(0.912)	Religious	0.572	0.451*
College grad	8.645	1.763		(0.292)	(0.174)
	(10.334)	(1.029)	Constant	0.017*	0.185
Fair/poor health	0.357	0.928		(0.030)	(0.240)
	(0.390)	(0.527)	ln(δ)	2.36	3.463***
Need for cognition-enjoyment	1.398	1.034		(1.137)	(1.192)
	(0.525)	(0.264)			
Number Series	0.69	1.029	Pseudo R-squared	0.169	0.055
	(0.286)	(0.298)	p-value	0.043	0.683
Verbal Analogies	1.377	1.002	Observations	623	644
	(0.411)	(0.222)			

Note: Results from negative binomial regression with dependent variable number of missing responses out of set of nine questions that are asked using identical wording and format on both mail and web modes. Column (1) includes mail surveys from web-eligible mail invitees and submitted web surveys; column (2) also includes partial web surveys. Exponentiated coefficients (incidence rate ratios) reported. Standard errors if the incidence rate ratios in parentheses. The data sources are the Cognitive Economics Study and the Cognition and Aging in the USA Study. * p<0.10, ** p<0.05, *** p<0.01

Appendix Table 3: Logit regression of response status, including complete, partial and switched mode submissions (average marginal effects of covariates omitted in Table 4)

	(1)	(2)	(3)		(1)	(2)	(3)
Age 50–54		-0.021	-0.015	WJ III: Verbal Analogies			-0.018
		(0.065)	(0.068)				(0.015)
Age 55–59		0.031	0.045	Low mental status			-0.036
		(0.064)	(0.066)				(0.038)
Age 60–64		-0.027	-0.022	Big 5 Personality: Extroversion			-0.02
		(0.067)	(0.070)				(0.014)
Age 65–69		0.066	0.079	Big 5 Personality: Agreeableness			0.01
		(0.067)	(0.069)				(0.014)
Age 70–74		0.032	0.035	Big 5 Personality: Conscientiousness			-0.01
		(0.072)	(0.075)				(0.014)
Age 75+		0.083	0.101	Big 5 Personality: Neuroticism			-0.002
		(0.066)	(0.068)				(0.015)
Female		-0.018	-0.003	Big 5 Personality: Openness			-0.021
		(0.026)	(0.032)				(0.017)
Coupled		0.036	0.043	Military			0.007
		(0.029)	(0.030)				(0.043)
Some College		-0.017	-0.025	Liberal			0.028
		(0.036)	(0.037)				(0.033)
College Graduate		0.021	-0.002	Conservative			0.017
		(0.035)	(0.039)				(0.029)
Fair/Poor Health		-0.063	-0.072*	Religious			-0.022
		(0.034)	(0.035)				(0.026)
Need for Cognition: Enjoyment			0.041*	Pseudo R-Squared	0.015	0.039	0.061
			(0.017)	P-value	0.001	0.007	0.025
WJ III: Number Series			0.004	Observations	813	813	793
			(0.020)				

Note: Results from logit regressions with dependent variable that is one if the respondent submitted a complete survey in the assigned mode, a partially complete web survey, or switched modes and submitted a survey, and zero otherwise. The data sources are the Cognitive Economics Study and the Cognition and Aging in the USA Study. * for $p<.05$, ** for $p<.01$, and *** for $p<.001$

	(1)	(2)	(3)	(4)
Outcome	Item non-response		Respond with exact value	
Age 50–54	-0.002	0.032	0.005	0.006
	(0.024)	(0.027)	(0.053)	(0.052)
Age 55–59	-0.001	0.003	0.04	0.041
	(0.024)	(0.023)	(0.052)	(0.052)
Age 60–64	-0.022	-0.021	0.005	0.005
	(0.023)	(0.021)	(0.050)	(0.050)
Age 65–69	0.035	0.075	0.013	0.014
	(0.034)	(0.039)	(0.050)	(0.050)
Age 70–74	0.011	0.037	0.007	-0.004
	(0.036)	(0.040)	(0.050)	(0.051)
Age 75 +	-0.006	0.012	0.029	0.025
	(0.027)	(0.029)	(0.050)	(0.050)
Female	0.021	0.021	0.004	0.003
	(0.017)	(0.020)	(0.021)	(0.021)
Coupled	0.013	-0.009	0.03	0.031
	(0.016)	(0.018)	(0.020)	(0.019)
Some college	-0.01	-0.022	0.014	0.011
	(0.016)	(0.021)	(0.024)	(0.024)
College grad	0.02	0.003	0.009	0.006
	(0.016)	(0.021)	(0.022)	(0.022)
Fair/poor health	-0.005	0.006	-0.012	-0.011
	(0.016)	(0.021)	(0.019)	(0.018)
Need for cognition-enjoyment	-0.001	0.007	0.008	0.009
	(0.008)	(0.010)	(0.011)	(0.011)
Number Series	0.009	0.016	0.019	0.019
	(0.009)	(0.010)	(0.011)	(0.011)
Verbal Analogies	-0.005	-0.014	-0.013	-0.014
	(0.008)	(0.009)	(0.008)	(0.008)
Low mental status	0.044**	0.024	0.011	0.014
	(0.014)	(0.019)	(0.024)	(0.024)
Personality: Extroversion	-0.005	-0.002	0.003	0.005
	(0.007)	(0.008)	(0.008)	(0.008)
Personality: Agreeableness	0.001	0.003	-0.004	-0.003
	(0.006)	(0.008)	(0.010)	(0.010)
Personality: Conscientiousness	0.008	0.014	0.001	0.001
	(0.007)	(0.010)	(0.008)	(0.008)
Personality: Neuroticism	0.007	0.008	0.009	0.01
	(0.006)	(0.008)	(0.009)	(0.009)
Personality: Openness	0.004	-0.006	0.004	0.003
	(0.008)	(0.011)	(0.009)	(0.009)
Military	0.005	-0.025	-0.039	-0.039
	(0.020)	(0.028)	(0.021)	(0.021)
Liberal	0.006	-0.025	-0.023	-0.027
	(0.014)	(0.019)	(0.021)	(0.021)
Conservative	0.027	0.023	-0.007	-0.009
	(0.015)	(0.017)	(0.018)	(0.018)
Religious	-0.015	-0.033*	0.007	0.01
	(0.012)	(0.016)	(0.016)	(0.016)
Pseudo R2	0.147	0.129	0.574	0.57
Observations	1869	1932	1800	1833