Effect of Alternative Data Collection Modes on Cooperation Rates and Data Quality

Sarah E. Brady, Courtney N. Stapleton, Julie A. Bouffard, Jerry D. Imel
U.S. Bureau of the Census, Washington, D.C. 20233

Key Words: Internet, Interactive Voice Response, Item Nonresponse

1. BACKGROUND[1]

The Census Bureau implemented the 2003 National Census Test (NCT) from January to April 2003. The objective of the test was tri-fold: to study the impact of offering various self-response options, new or additional contact strategies, and alternative race and Hispanic origin questions on cooperation rates and data quality. The overall goal of the 2003 NCT was to identify the best strategy for increasing self response to the census, thus reducing the Nonresponse Followup (NRFU) workload. Successful accomplishment of this goal would improve the data quality of the 2010 Census while reducing the cost of data collection.

This paper presents the methodology and results of the self-response option portion of the NCT. The self-response options tested the following response modes: Internet, Interactive Voice Response (IVR), and paper. The analysis presented in this paper assesses the effect of offering these alternative response modes and the interactions among these various modes on overall cooperation rates and data quality.

2. METHODOLOGY

2.1 Panel Design

The methodology for the 2003 NCT consisted of a data collection strategy involving sixteen different experimental panels. One panel represented a control group that received a four-component mailing strategy including an advance letter, initial questionnaire package, reminder postcard, and replacement questionnaire (sent to nonresponding households only). This panel was the control for all three portions of the 2003 NCT.

The self-response option portion of the test comprised five panels. These five panels examined strategies for increasing self response to the census using alternative response modes. Specifically, the panels tested the impact of offering different combinations of paper, Internet, and IVR reporting options.

The electronic response options reflected two strategies referred to as 'push' and 'choice'. Households in the push panels did not initially receive a paper questionnaire. In place of a paper questionnaire, those households initially received a guide for using one or both electronic response options. The letter sent with the guide requested that households use one of the electronic response options. Households in the choice panels were offered the electronic response options in addition to the option of mailing back a paper questionnaire, with no language pushing the respondent to use an electronic mode. Households assigned to any of the panels that included the IVR option were not told that it was an automated system.

Specifically, the 2003 NCT self-response option panels were:

Control: Households received an advance letter, an initial questionnaire, and a reminder postcard. Nonrespondents received a replacement questionnaire (RQ).

Similar to the Control panel, all of the response mode panels included an advance letter, a reminder postcard, and a replacement questionnaire sent to nonrespondents.

Push IVR (M1): Households initially received a guide to the IVR system in place of a paper questionnaire. This panel tested a push strategy for the IVR system.

IVR Choice (M2): Households received an initial paper questionnaire that included information about the IVR system. Households could choose to respond via paper or the IVR system. This panel tested a choice strategy for the IVR system.

Internet Choice (M3): Households received an initial paper questionnaire that included information about the Internet application. Households could choose to respond via paper or the Internet. This panel tested a choice strategy for the Internet.

Push IVR and Internet (M4): Households initially received a guide to both electronic response options in place of a paper questionnaire. This panel tested a push strategy for the IVR and Internet.

IVR and Internet Choice (M5): Households received an initial paper questionnaire that included information about both electronic response options. Households could choose to respond via paper, IVR, or Internet. The M5 panel tested a choice strategy for the IVR and Internet.

2.2 Mailing Strategy

The mailing strategy for the self-response option panels used a multiple contact approach. The content of each contact was dependent upon the panel assignment. The mailing strategy consisted of four mailings. In general, these mailings were as follows: an advance letter for the first mailing; an initial request to complete census questions as the second mailing; a reminder postcard as the third mailing; and a replacement questionnaire for nonrespondents as the fourth and final mailing.

The second mailing was where the mailing strategy differed across panel. Households in the Control panel or one of the

[1] This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress.

three choice panels received a paper questionnaire. In addition, questionnaire packages sent to those in the choice panels included brief instructions for the electronic response options – either the website address, IVR telephone number or both, depending on panel assignment. Households in the two push panels received a guide to completing their census form electronically (either by Internet or IVR) in lieu of a paper questionnaire. Regardless of panel, the mailing package included a letter from the Census Bureau's Director urging households to respond. The letter for the push panels also used motivational language about the IVR and/or Internet systems to explain the benefits of the electronic systems, such as ease, convenience, and cost savings to taxpayers.

For the fourth and final mailing (the targeted replacement questionnaire), the replacement questionnaire was sent to all housing units that had not responded prior to February 12th. The previous mailings did not inform the respondents that they would receive a paper replacement questionnaire if they did not return their initial questionnaire. Both the Control and choice panels received the same set of materials as in the initial mailing package. However, at this stage, households in the push panels were provided with a paper replacement questionnaire, but were still encouraged to respond via their originally assigned electronic mode(s). Thus, for push panels, only nonrespondents ever received a paper questionnaire.

2.3 Modes

Depending on their panel assignment, housing units selected for the 2003 NCT were eligible to respond by up to three modes. The three modes were paper, Internet, and IVR. Each data collection mode collected the census short form data items.

2.3.1 Paper

The 2003 NCT form was modeled after the Census 2000 short form, with only minor changes in the introductory language (to reflect the 'test' nature of the form). The form allowed the respondent to list names for up to 12 household members. For up to six household members, the form provided space for reporting the 100 percent census data items (i.e. name, relationship, sex, age/date of birth (DOB), Hispanic origin, and race). The form also collected traditional short form housing unit level data (household count, home ownership, and telephone number).

2.3.2 Internet

The 2003 NCT Internet application collected the same data as the paper form. For the 2003 NCT, we had the constraint that the wording and flow of the census questions in the Internet application had to mirror the format from the paper questionnaire. This decision was made to minimize the possibility that the cooperation rates were due to questionnaire changes, rather than the offering of a specific mode. The Internet application required respondents to enter their 14-digit Housing Unit ID (HUID), which was printed on the materials they received in the mail. The application collected short form housing unit level and person level data for up to 30 household members. The interactive application included a progress indicator on the left-hand side of the screen and allowed respondents to back up and correct previously entered information. The system incorporated soft edits, which alerted respondents to incomplete or inaccurate responses but did not require corrections to these items. Once respondents entered data for all household members, they could view a summary of their answers prior to making a final submission.

2.3.3 Interactive Voice Response (IVR)

As with the Internet application, the 2003 NCT IVR application collected the same data as the paper form and the wording and flow of the census questions in the IVR application had to mirror the format from the paper questionnaire. This constraint resulted in dialogue that was not optimal for an IVR application.

The IVR application used speech recognition technology. That is, respondents were asked to speak their answers, and the system detected their responses by comparing them to a set of "in vocabulary" responses for the questions. The IVR system then repeated back to the respondent what the system "understood" for verification. However, in some cases the respondents may not have answered the question. In addition, during verification, the respondents may have indicated that the IVR system heard the response incorrectly. In these two instances the IVR system re-prompted the respondents with a slightly altered wording of the question. The altered wording either provided more information or let the respondents know that they could use touchtone entry to key in their response. However, if the system was still unable to understand the respondent after the maximum number of re-prompts (in general, two) then that question was considered a failure. If the question was one that the Census Bureau had determined as critical for further census processing, such as HUID, then the respondent failed out of the IVR system and the call transferred to a telephone agent. Similarly, if there were two consecutive question failures, the call transferred to a telephone agent. Otherwise, the system moved on to the next question.

When a call was transferred, the agent obtained the respondent's census information by conducting an interview from the beginning of the survey (data from the IVR did not carry over into the agent interview) and submitted the data via a modified version of the Internet application.

2.4 Sample Design

We selected seven panels of 10,000 housing units each and nine panels of 20,000 housing units each, for a total of 250,000 housing units to form sixteen panels for the 2003 NCT. (Due to the proposed item level analysis, some panels required the larger sample size of 20,000 housing units.) For purposes of this paper, only the response mode panels and the Control panel will be discussed. The self-response mode portion of the test consisted of 70,000 housing units; the Control panel contained 20,000 housing units and all response mode panels contained 10,000 housing units each. Each of the response mode panels' sample size was divided evenly into

two strata, the low response area stratum and the high response area stratum.

Prior to sample selection, census tracts were stratified into two groups that reflect differences in Census 2000 mail return rates, as well as anticipated differences in the race/Hispanic origin and home ownership composition (owner vs. renter-occupied housing units) of the population. For the High Response Area (HRA) stratum, the average mail return rate was 81.4 percent. For the Low Response Area (LRA) stratum, the average mail return rate was 62.2 percent.

The LRA stratum was expected to contain a very high proportion of the Black and Hispanic populations and renter-occupied housing units. The addresses in the LRA stratum were sampled at a higher rate than those in the HRA stratum, to ensure sufficient representation of the low response areas. Estimates presented in this paper were weighted to account for oversampling of the LRA stratum.

2.5 Calculation of Cooperation Rates

Cooperation rates[2] are a measure of respondent behavior with regard to returning a questionnaire. It is defined as the number of primary returns divided by the number of sample cases in the panel less the number of ineligible sample cases, which were cases returned by the United States Postal Service (USPS) as Undeliverable as Addressed (UAA) for that panel.

$$Cooperation\ Rate = \frac{\#of\ primary\ returns}{Panel\ sample\ size - ineligible units\ for\ the\ panel}$$

2.5.1 Cooperation Rate Denominator

The denominator for the cooperation rate was all sample cases for the panel after removing those cases that were determined to be UAA. UAAs were defined on a housing unit basis as a unit having any of three mailing pieces (initial questionnaire, reminder postcard, or replacement questionnaire) returned by the USPS. Any housing units determined to be UAA was considered an ineligible unit.

Housing units classified as UAA were removed from the denominator and any returns received for UAA housing units were not included in the numerator. Due to the way we classified a housing unit as UAA, a return could be received for a housing unit that was classified as UAA. For example, the initial questionnaire was successfully delivered and a household member completed and returned the questionnaire. However, the reminder postcard was returned by the USPS as UAA, which resulted in the housing unit being classified as UAA. We observed an approximate 10 percent UAA rate for each panel.

[2] Our definition of cooperation rate is in line with the definition established by the American Association of Public Opinion Research. (However, there is a slight difference in how we determine eligible units. See Section 2.5.1.)

2.5.2 Cooperation Rate Numerator

The numerator for the cooperation rate was defined by all returns for the panel after removing those cases that were determined to be blank or duplicates. In addition, there were returns removed from the numerator because the corresponding housing unit was classified as UAA.

2.5.2.1 Blank Returns

The blank return definition was restricted to Persons 1-6. We defined a blank return as an eligible return with fewer than two "completed" census data items. The data items examined for completeness were: name, relationship, sex, age or date of birth, Hispanic origin, race, home ownership, and household count. The percent of blank returns with valid HUIDs across all panels ranged from 1.5 to 4.7 percent.

2.5.2.2 Duplicate Returns

If duplicate returns were received for a given housing unit, we accepted the first nonblank return for data analysis. There was one exception to this rule: when the first nonblank return was an IVR return that had failed out of the IVR system and was flagged as a transfer to an agent. For these cases, we selected an agent return or the next non-blank return received. If none of these existed, we selected the IVR failure return.

2.6 Calculation of Item Nonresponse Rates

Item nonresponse rates are a measure of the 2003 NCT data quality. It refers to the proportion of records with missing data for a particular item. This analysis was restricted to nonblank, primary returns for a housing unit. Item nonresponse rates were calculated according to the following definition:

$$Item\ Nonresponse\ Rate = \frac{\#\ of\ records\ with\ missing\ data\ for\ a\ particular\ item}{total\ number\ of\ records}$$

Item nonresponse rates were calculated for both housing unit level items and person level items. So, for the housing unit level items, the term 'records' in the item nonresponse rate definition refers to housing units. The total number of housing units was defined as the number of housing units from all nonblank primary returns. For person level items, the term 'records' refers to persons. The total number of persons was defined as the number of persons listed on all nonblank primary returns.

For this analysis, we calculated item nonresponse rates for five person level data items (relationship, sex, age/year of birth (YOB), Hispanic origin, and race) and two housing unit level data items (home ownership and household count). In order to determine missing values for data items, we looked at the presence or absence of a value for a particular item, not whether that item's value was consistent with other data items reported.

2.7 Variance Estimation

To account for the stratified clustered sample design, WesVarPC version 4.1 was used to compute standard errors for all estimates. A jackknife replication methodology using random groups was used to estimate standard errors. The housing units were sorted in the same order that they were selected and the clusters of housing units (or housing units selected at each hit) were assigned sequentially to one of the 250 random groups.

3. LIMITATION

As part of the security plan for the IVR application, log files were to be deleted after a period of seven days. The first time the deletion was performed, a 'bug' was discovered in the software used to delete the log files. This bug caused some necessary links in the application to be deleted, which resulted in respondents receiving an error message at a point near the beginning of the IVR instrument. The links were restored approximately 38 hours later.

The exact behavior of the IVR system during this period is unclear. Test calls to the IVR system illustrated that the behavior of the IVR instrument was inconsistent during this time. It was determined that some calls with households of size one were able to complete the interview without getting the error message. However, households of size two or more always got an error during this down time. Analysis showed that any call that got the error message should have transferred to a telephone agent. However, corresponding telephone agent records were not found for all transferred calls during this time.

This problem has item nonresponse implications since these records had the potential to be selected as the primary return for the household. As a result of this problem, IVR returns associated with the down period were excluded from the item nonresponse analysis. Note that these IVR returns were not dropped from the cooperation rate analysis. Also note we did not drop any agent returns that were completed during the down period.

4. RESULTS

4.1 Cooperation Rate Analysis

Table 1 contains the cooperation rates at the national level.

Table 1. Cooperation Rates at the National Level

	Estimate	S.E.
Control	67.1%	0.39%
Push IVR (M1)	62.2%	0.67%
IVR Choice (M2)	65.9%	0.58%
Internet Choice (M3)	67.3%	0.62%
Push IVR & Internet (M4)	61.4%	0.70%
IVR & Internet Choice (M5)	66.4%	0.66%

Table 1 illustrates that the choice panels (M2, M3, and M5) showed cooperation rates comparable to the Control panel. The push panels (M1 and M4), however, showed lower cooperation rates.

When we looked at the results by strata, we saw that, in general, cooperation rates from the high response area were about 20 percentage points higher than low response area rates. This was expected based on the sample design.

In order to determine the effect the alternative modes had on the cooperation rates, differences in the cooperation rates were calculated between each of the alternative mode panels and the Control panel. These differences are in Table 2.

Table 2. Comparisons of the Alternative Mode Panels with the Control Panel at the National Level

	Difference	90 % C.I.
M1 – Control	-4.9*	-6.61 to -3.19
M2 – Control	-1.2	-2.80 to 0.40
M3 – Control	0.2	-1.40 to 1.80
M4 – Control	-5.7*	-7.50 to -3.90
M5 – Control	-0.7	-2.41 to 1.01

**Significant at alpha = 0.10 and critical value = 2.28*

As seen in Table 2, both push panels (M1 and M4) had significantly lower cooperation rates than the Control panel. That is, offering IVR or IVR & Internet with a push strategy significantly decreases cooperation rates, 4.9 percentage points and 5.7 percentage points, respectively. Push panels had significantly lower cooperation rates than the Control panel for the high and low response strata.

The three choice panels (M2, M3, and M5) showed no significant differences from the Control panel, nationally and by both high and low response strata. That is, offering a choice of response modes shifts response across modes, rather than increasing overall response. See Section 4.2 for more details. This finding confirms the results of the 1993 Mail and Telephone Mode Test, which tested offering a telephone response option in addition to paper, and found only a shift in response across modes, rather than an increase in overall response (U.S. Bureau of the Census, 1993).

Table 3. Alternative Mode Panel Comparisons (Push Panels vs. Choice Panels) at the National Level

	Difference	90% C.I.
M1 – M2	-3.7*	-5.07 to -2.33
M4 – M5	-5.1*	-6.75 to -3.45

** Significant at alpha = 0.10 and critical value = 1.96*

As seen in Table 3, the Push IVR (M1) panel had significantly lower cooperation rates than the IVR Choice (M2) panel. This result was also observed for both high and low response strata. That is, households who were initially given the choice between IVR and paper were more likely to respond than households initially asked to respond by IVR only.

The Push IVR & Internet (M4) panel had a significantly lower cooperation rate than the IVR & Internet Choice (M5) panel, nationally and by both high and low response strata. That is,

households who were initially given the choice between IVR, Internet, and paper were more likely to respond than households initially asked to respond by IVR or Internet only.

4.2 Distribution of Response by Mode

The distribution of overall response for each mode was examined, that is how much of the response for each panel was due to a specific mode. The results of this analysis are presented in Table 4, which is located at the end of this report.

Table 4 illustrates that the push strategy succeeded in pushing respondents to an electronic mode. In the Push IVR panel (M1), over 67 percent of respondents used the IVR system. In the Push IVR & Internet panel (M4), over 38 percent used the IVR system and over 30 percent responded by Internet. However, the push panels had significantly lower national cooperation rates (when compared to the control) as shown in Table 2.

When given a choice between using an electronic response mode and paper (M2, M3, and M5 panels), respondents largely chose to respond by paper (88.4 percent to 93.2 percent). However, a slight shift to the electronic modes did occur, up to 11.6 percentage points for the M5 panel. That is, offering alternative response modes did not increase overall response. Instead, many who would have otherwise responded by paper shifted to IVR or Internet.

Across all panels that included the IVR option, 17 to 22 percent of IVR primary returns were ultimately completed by telephone agents. Considering the size and scope of a decennial census, the infrastructure required to accommodate this number of calls could be very extensive. A reduction in the proportion of calls transferred to agents would certainly be required. The 17 to 22 percent also suggests that a good number of IVR respondents had difficulty with the IVR system. See Section 4.4 for a more detailed discussion of these difficulties.

Some households assigned to the M1 and M2 panels (where IVR is the only electronic mode offered) managed to respond via the Internet. However, the effect on the distribution of response by mode due to this is minimal.

4.3 Timing of Returns by Mode

The percent of returns received were calculated by mode for the Control panel and the M5 panel for each day of the 2003 NCT, beginning with January 30, the start of the initial questionnaire delivery. Figure 1 displays the timing of returns for the Control and M5 panels by mode. Note that paper forms were not checked in on weekends and holidays, therefore cooperation rates did not increase for those corresponding days. This is reflected in the observed spikes in response on day 12, day 20, and day 26.

Figure 1. Timing of Returns for Control and M5 Panels by Mode

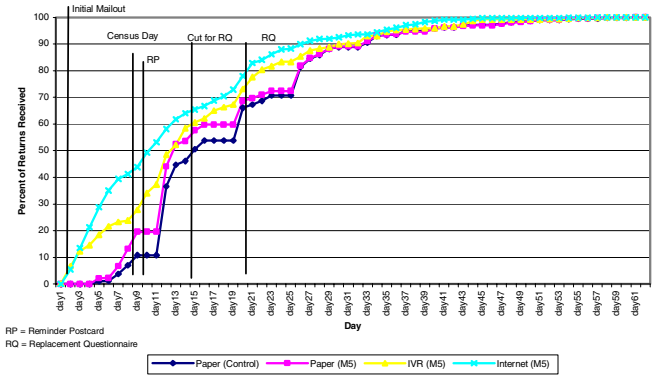


Figure 1 illustrates that returns submitted by electronic modes were received at a quicker pace than paper. As of Census Day, approximately 23 percent of the IVR returns and 40 percent of the Internet returns had been received, compared to only about 10 percent of the paper returns (both Control and M5 panels). This points out a key advantage of offering electronic modes. When given the choice of paper, Internet, and IVR, the Internet responses come in the fastest. IVR responses come in almost as quickly, and paper forms lag behind before ultimately catching up. Speed is an important issue in the context of a decennial census, especially when considering a targeted replacement questionnaire. The earlier returns are received, the fewer targeted replacement questionnaires sent out.

We also see that paper respondents in the M5 panel function differently than the Control group respondents. Paper respondents in the M5 panel responded sooner than paper respondents in the Control group. We believe this to be an effect of offering the electronic modes. That is, offering more 'immediate' response modes may have triggered respondents to respond sooner, even if they did not choose one of the electronic modes. Electronic modes may indicate to respondents that their response is urgent, or perhaps give respondents the sense that the Census Bureau is cooperative and/or considerate to offer multiple response options, thus encouraging response sooner.

4.4 Item Nonresponse Rate Analysis

Table 5, which is located at the end of this paper, contains the national level item nonresponse rates by mode.

Table 5 shows item nonresponse rates for person and housing unit level items by mode for all persons. Initial analysis results indicated that, in general, there were no significant differences in item nonresponse rates by panel for each of the modes, therefore we combined treatment panels by mode. Results were also analyzed at the stratum level, and as expected, item nonresponse was higher in the low response stratum, and lower in the high response stratum.

Table 5 shows that for nearly all Internet items at the person level, item nonresponse rates were significantly lower when compared to paper returns in the Control panel. Specifically, Control paper item nonresponse rates for Hispanic origin and race were more than twice the Internet item nonresponse rates. The Hispanic origin and race items were on two separate screens in the Internet application and the application did not include the instructions to answer both items as compared to the paper form. However, we do not attribute the lower item nonresponse rates for the race and Hispanic origin items to this design. The lower item nonresponse rates for the Internet could have been due in part to the design of the application. In the Internet application, response to the data items was not required, but respondents were given a one-time edit if they failed to initially respond to the question.

For all person level data items, we observe significantly higher item nonresponse rates for the IVR system as compared to paper returns in the Control panel. IVR item nonresponse rates for sex, age, Hispanic origin, and race range from eight to nine percent. For the most part, we don't believe that the high IVR item nonresponse rates were a reflection of difficulty with the items themselves, but rather, difficulty with the implementation of the IVR system. The high IVR rates were also somewhat due to the way in which blank person records were created for this mode. That is, once the respondent entered the number of persons in the household, the IVR system created all of the person records for the household. Therefore, if the respondent hung up or transferred during the course of the interview, all remaining items for all remaining household members were considered missing. However for the Internet and paper returns, in general, person records were created only after the respondent began entering data for that person. Therefore, due to this difference in the creation of person records, IVR item nonresponse rates for person level data items may be slightly inflated.

The IVR item nonresponse rate for the relationship item was even higher, at 12.8 percent. We believe this rate to be a reflection of the implementation of the relationship question in the IVR system. The relationship question in the IVR system involved a branching technique. Therefore, before hearing all of the individual relationship categories, respondents first had to select 'Relative', 'Roomer', 'Boarder', 'Housemate', 'Roommate', Unmarried Partner', 'Foster Child', or 'Other Non-relative'. Immediate family members such as spouses, children and parents were listed under the 'Relative' category. We believe the concept of classifying immediate family members as 'relatives' was difficult for some respondents. In addition, the response options presented to the respondents were probably not what they expected to hear. If respondents had looked at the IVR guide or paper census form, they probably expected to hear the categories as they were listed on the form/guide which included the full list of categories for the relationship question.

The IVR system did not affect the data quality for transferred cases. Agent item nonresponse rates for person level data were very low. That is, when cases rolled over from the IVR system, respondents provided just as good or better data, as measured by item nonresponse rates, than had they filled out a paper form in the Control panel. This reflected a great deal of patience on the part of the respondents, given that they had to start over from the beginning of the census form once they transferred to an agent.

At the housing unit level for IVR, Internet, and agent, we observe significantly lower item nonresponse for the housing unit level items (home ownership and household count) as compared to paper returns in the Control. However, it should be noted that these rates were a reflection of the requirements associated with the electronic modes. That is, IVR respondents were required to answer the household count question in order to continue through the application. As previously stated, in the Internet application, response to the housing unit level items was not required, but respondents were given a one-time edit if they failed to initially respond to the questions.

The item nonresponse rate for household count for the Control paper returns was more than twice that of the corresponding rate for the paper returns associated with the alternative panels (Paper M1-M5). We further analyzed the paper returns from the alternative panels, and found no differences among push and choice panels. However, we discovered that the design of the paper form for the alternative mode panels was slightly different than the Control panel paper form. That is, we included the web address and/or the IVR telephone number, depending on panel, in a squared off section at the top of the form. This section directly preceded the household count question. Therefore, we speculate that this design feature brought respondent's attention toward the household count question, making them more likely to notice, and answer, that question. Examples of the two forms can be found in Figures 2 and 3 at the end of the report.

5. CONCLUSIONS AND RECOMMENDATIONS

The results of the alternative mode portion of the 2003 NCT illustrated that:

- Offering a choice of alternative modes of response did not increase or decrease the cooperation rates. Instead it simply shifted response to the alternative modes. However, this shift was relatively small.

- Pushing respondents to respond by electronic modes was found to decrease overall response. However, more analysis of the push strategies should be conducted, as at least two-thirds of the response for push panels came in via an electronic mode, indicating that a push strategy may have potential for future use.

- Choice panels had significantly higher cooperation rates than the push panels. Households who were initially given the choice between IVR and paper were more likely to respond than households initially pushed to respond by IVR. In addition, households who were initially given the choice between IVR, Internet, and paper were more likely to respond than households initially pushed to respond by IVR or Internet.

- Returns submitted by electronic modes were received at a quicker pace than paper. This points out a key advantage of offering electronic modes. When given the choice of paper, Internet, and IVR, the Internet responses come in the fastest. IVR returns come in almost as quickly, and paper forms lag behind before ultimately catching up. Speed is an important issue in the context of a decennial census, especially when considering a targeted replacement questionnaire. The earlier forms are received, the fewer targeted replacement questionnaires sent out.

- For all person level items, there were significantly higher item nonresponse rates for the IVR returns than for paper returns associated with the Control panel. These rates were likely a reflection of the difficulty with the IVR system rather than difficulty with the items themselves. The item nonresponse rates for all housing unit level items were significantly lower for IVR returns than for paper returns from the Control panel. The lower rates for IVR housing unit level items were a result of the requirements associated with the system.

- The quality of data from Internet returns was significantly better than the quality of data associated with paper returns, as measured by item nonresponse rates. For nearly all Internet items, the item nonresponse rate was significantly lower for Internet returns than for rates associated with paper returns from the Control panel. These lower rates could have been due to the soft edits in the Internet application.

Based on these conclusions we recommend the following:

- **Continue experiments with electronic data collection systems-** As we move into the future, respondents are likely to become accustomed to these alternative modes and usage may increase.

- **Develop content suitable for each mode-** The challenge we face is to optimize the advantages of each mode while collecting the same data across modes. We need to first develop a better instrument for each mode, while ensuring that each mode obtains the same information. We would then test each mode against the other modes to verify we are collecting the same data. Future research should include a content reinterview to get at variation across modes. Furthermore, we should research the design of the instruments so as not to compromise data quality, while maximizing the advantages of each mode.

- **Research improvements to the IVR system-** We recognize the respondent burden associated with using the IVR during this test. That is, we saw high item nonresponse rates and a large number of transfers to agents. Research should continue on the IVR as a survey response mode, and continuous testing should occur to develop the best system possible to decrease item nonresponse rates and decrease agent workloads.

- **Timing of offering alternative modes-** We should continue research on the most appropriate timing of offering alternative modes, specifically in terms of push and choice strategies.

- **Box design above household count question-** We saw the item nonresponse rate for the household count question was significantly lower for the paper returns from the alternative mode panels as compared to the rate for the paper returns from the Control panel. Upon further investigation, we determined that this result was due to the forms design for the alternative mode panels. That is, when we added the alternative mode information to the initial questionnaire a line was used to separate this information from the first question on the form, the household count question, forming a box-like design. We concluded that this drew attention to the question, as reflected by the increased response to that item for the paper returns from the alternative mode panels. We recommend using a similar box design for future questionnaires.

ACKNOWLEDGEMENTS

We would like to acknowledge the assistance of the following people: Mike Bentley, Joan M. Hill, Tracy Mattingly, Dave Sheppard, Jim Treat, and Henry Woltman.

REFERENCES

U.S. Bureau of the Census, 1993, *Final Results of the Mail and Telephone Mode Test,* Decennial Statistical Studies Division 2000 Census Memorandum Series #E-47, July 21, 1993.

U.S. Bureau of the Census, 1999, *Definition of a Blank Form by Data Capture System 2000*, Decennial Statistical Studies Division Census 2000 Memo Series # K-3, March 8, 1999.

U.S. Bureau of the Census, 2002a, *Program Master Plan for the 2003 National Census Test*, 2003 National Census Test (NCT) #02-05, July 2002.

U.S. Bureau of the Census, 2002b, *Sampling Specifications for the 2003 National Census Test*, Planning, Research and Evaluation Division TXE/2010 Memorandum Series #TX-RASE-S-3, September 16, 2002.

U.S. Bureau of the Census, 2003, *2003 National Census Test Analysis Plan (DRAFT),* March 4, 2003.

Table 4. Distribution of Response by Panel and Mode

Panel	Paper n=29,938		Internet n=2,457		IVR** n=5,999		Agent*** n=1,179	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Control	100.0%	0.00%	--	--	--	--	--	--
Push IVR (M1)	32.3%	0.75%	0.6%	0.11%	67.1%	0.73%	17.2%	0.65%
IVR Choice (M2)	93.2%	0.36%	***0.0%****	***0.01***	6.8%	0.36%	17.9%	2.22%
Internet Choice (M3)	90.0%	0.42%	10.0%	0.42%	--	--	--	--
Push IVR & Internet (M4)	31.1%	0.74%	30.7%	0.77%	38.2%	0.74%	20.6%	0.99%
IVR & Internet Choice (M5)	88.4%	0.45%	7.1%	0.37%	4.5%	0.29%	21.9%	2.55%

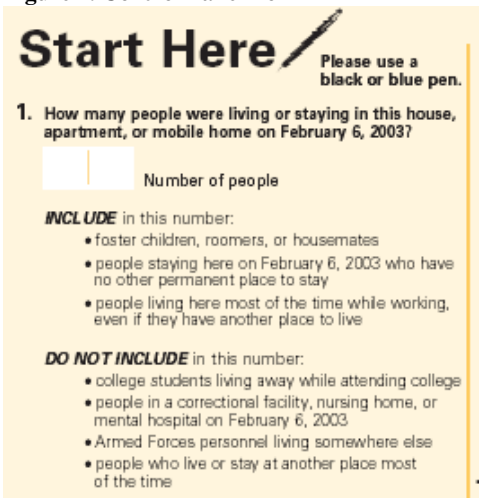
Numbers presented in **bold italics are based on cell sizes less than 10 and should not be used for comparison.*
***IVR column includes IVR and Agent primary returns*
****Agent column represents Agent values as a percent of IVR column*

Table 5. Item Nonresponse Rates by Mode, National Level

Item	Paper (Control)	Paper (M1-M5)	IVR	Internet	Agent
	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)
Relationship	2.1% (0.17%)	2.4% (0.16%)	12.8%* (0.85%)	0.5%* (0.13%)	1.9% (0.40%)
Sex	1.8% (0.12%)	1.9% (0.10%)	8.0%* (0.55%)	1.2%* (0.15%)	1.5% (0.28%)
Age/YOB	1.7% (0.12%)	1.9% (0.11%)	8.4%* (0.55%)	1.5% (0.24%)	1.5% (0.31%)
Hispanic Origin	3.8% (0.17%)	4.0% (0.15%)	9.0%* (0.56%)	1.5%* (0.16%)	1.4%* (0.26%)
Race	3.8% (0.20%)	4.1% (0.16%)	8.7%* (0.58%)	1.5%* (0.17%)	1.5%* (0.28%)
Home ownership	2.0% (0.14%)	1.6%* (0.09%)	0.8%* (0.13%)	0.2%* (0.09%)	0.4%* (0.21%)
Household count	2.7% (0.17%)	1.1%* (0.08%)	0.0%* (0.0%)	0.0%* (0.0%)	0.0%* (0.0%)

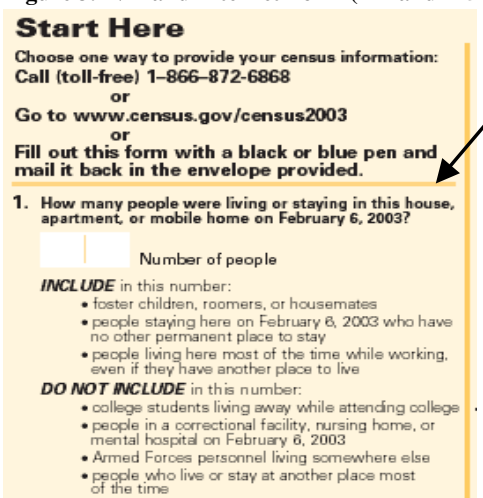
†Relationship excludes Person 1.
††IVR rates are slightly inflated due to how person records were created, see Section 4.4.
**Significantly different from Paper (Control) at alpha=0.10 and critical value=2.2*

Figure 2. Control Panel Form


Start Here
Please use a black or blue pen.
1. How many people were living or staying in this house, apartment, or mobile home on February 6, 2003?
Number of people
INCLUDE in this number:
- foster children, roomers, or housemates
- people staying here on February 6, 2003 who have no other permanent place to stay
- people living here most of the time while working, even if they have another place to live

DO NOT INCLUDE in this number:
- college students living away while attending college
- people in a correctional facility, nursing home, or mental hospital on February 6, 2003
- Armed Forces personnel living somewhere else
- people who live or stay at another place most of the time

Figure 3. IVR and Internet Form (M4 and M5 Panels)


Start Here
Choose one way to provide your census information:
Call (toll-free) 1-866-872-6868
or
Go to www.census.gov/census2003
or
Fill out this form with a black or blue pen and mail it back in the envelope provided.
1. How many people were living or staying in this house, apartment, or mobile home on February 6, 2003?
Number of people
INCLUDE in this number:
- foster children, roomers, or housemates
- people staying here on February 6, 2003 who have no other permanent place to stay
- people living here most of the time while working, even if they have another place to live

DO NOT INCLUDE in this number:
- college students living away while attending college
- people in a correctional facility, nursing home, or mental hospital on February 6, 2003
- Armed Forces personnel living somewhere else
- people who live or stay at another place most of the time

Line added to forms for alternative mode panels to separate mode information. Created box design on form above household count question.