NORC at the UNIVERSITY OF CHICAGO
Innovative solutions in research and technology

Analysis Plan

2010 Census Integrated Communications Program Evaluation (2010 CICPE)

National Opinion Research Center (NORC)
at the University of Chicago
55 East Monroe Street
Suite 3000
Chicago, IL 60603
(312) 759-4000
(312) 759-4004 (Fax)

February 2, 2010

Table of Contents

2010 CICPE Analysis Plan

This analysis plan describes analyses planned for the 2010 Census Integrated Communications Program Evaluation (2010 CICPE) being conducted by the National Opinion Research Center (NORC) at the University of Chicago. The analysis plan is structured as an outline for the final report that we envision for this evaluation, with annotations for the expected contents of each chapter of the report.

There have been many changes to the design of this study, and some details are still unresolved (for example, the inclusion of confirmed awareness and mode-specific cueing items in the Wave 3 questionnaire, and the extent to which information can be meaningfully incorporated into analyses from such sources as the Integrated Partner Contact Database (IPCD) and the VOCUS data base on earned media). Thus we consider this a draft which will evolve as the evaluation progresses and data analysis begins.

Chapter 1.
2010 Census Integrated Communications Program Evaluation
(2010 CICPE)

A. Background of 2010 CICPE

This section of the 2010 CICPE final report will discuss the history of paid media in decennial censuses, the objectives of the 2010 Integrated Communications Campaign (ICC), and its overall architecture.

B. Structural Model for this Evaluation

The structural model developed for this evaluation is shown on page 5 (below). This section of the final report will elaborate on the model, including discussion of how it is implemented in the analyses reported.

A further discussion will discuss the nature of the 2010 ICC, and implications of the campaign structure for the structural model. For example, NORC has cautioned that the 2010 CICPE will not be able to distinguish between certain campaign elements, such as regional versus national partnerships. We also anticipate extensive difficulty in separating out partnership effects from paid media and earned media effects, especially because we anticipate that the general public (and therefore the 2010 CICPE survey respondents) will not be able to distinguish their awareness of and exposure to these program elements.

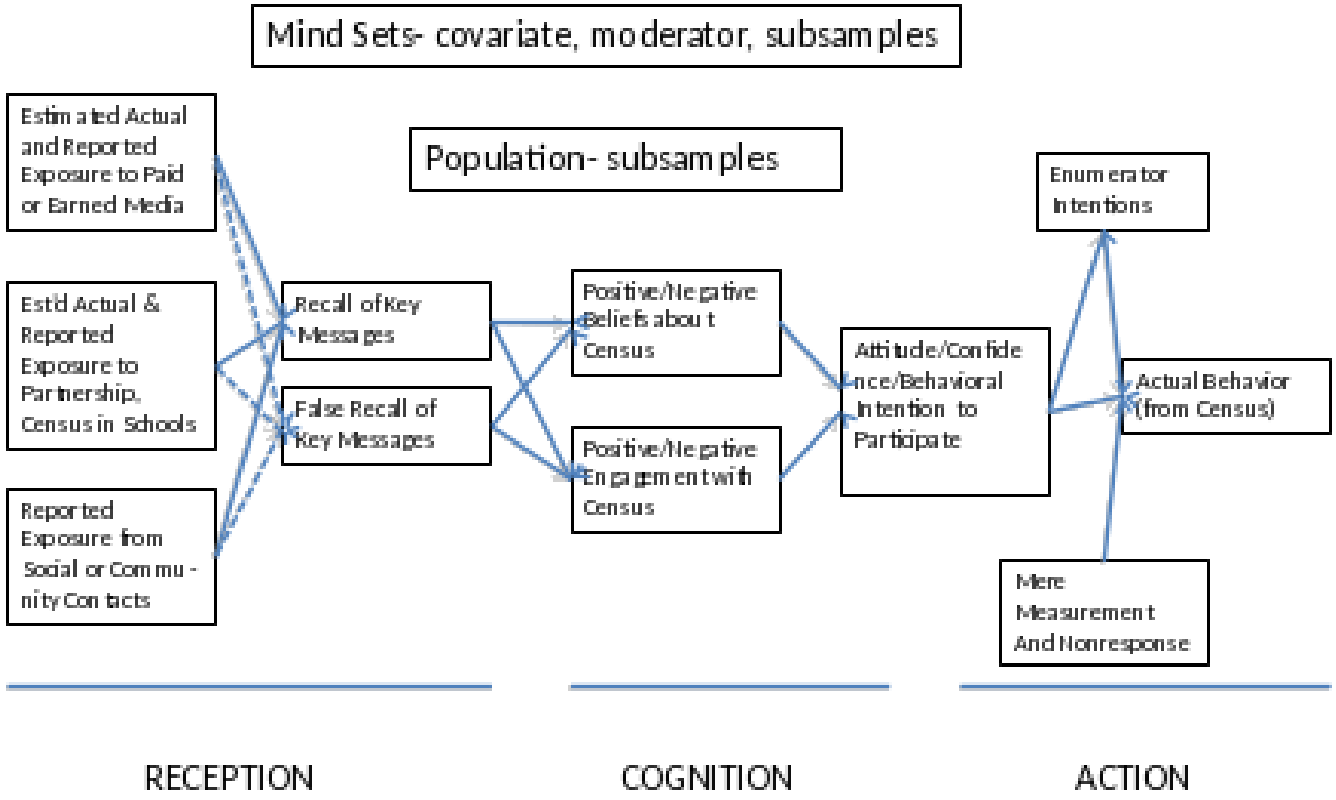
C. Evaluation Objectives

The objectives for this evaluation, have been articulated by the Census Bureau as follows: to assess the extent to which the 2010 ICC achieved a variety of specific goals related to **increased mail returns**, **improved accuracy through reduced differential undercount**, and **improved cooperation with enumerators**. Specific questions to be answered by the 2010 CICPE include:

1. How effective was the overall communications strategy at contributing to improvements in response accuracy, and cooperation with enumerators?
2. Which elements of the paid media/advertising were reported/recalled both least and most often?
3. Which elements of the Partnership Data Services (PDS) Program (National and Regional were reported/recalled both least often and most often?
4. How effective was the paid media/advertising campaign in changing positive and negative attitudes and beliefs about the Census?
5. How effective was the Partnership Data Services (PDS) Program in changing positive and negative attitudes and beliefs about the Census?

6. How effective was the “Census in Schools” Program in changing positive and negative attitudes and beliefs about the Census?
7. What impact did the 2010 Census Integrated Communication Campaign as a whole have on the likelihood of returning a Census form?
8. What differences in awareness, knowledge, and attitudes before during and after the 2010 Census Integrated Communication Campaign significantly different from those measured before, during, and after the 2000 Advertising Campaign?
9. What advertisements, programs and events (including breaking news events) outside of the 2010 integrated Communications Campaign had an effect on respondent attitudes and behaviors?
10. What return on investment can be estimated for the 2010 ICC?

Structural Model



D. Survey Design

In the final report this several-page section will essentially provide a methodology report on the design and data collection of each of three waves of survey data collection. The survey design explication will lay out the three survey waves (prior to paid media launch, during paid media campaign prior to census mailout, during non-response follow-up), the designs of the core (Black, Hispanic, and non-Black/non-Hispanic) and supplemental samples (Asian, NHOPI, AIAN), and the structure of the panel sample. Cases will be defined as panel sample cases based on their inclusion in the panel sample, whether or not they ever completed a second (or third) interview. Panel sample cases are selected after Wave 1 data collection for fielding in Waves 2 and 3.

In addition to such standard elements as data collection approach, sample design and response rates achieved, the section will include details relevant to the interpretation of data, for example, 1) respondent vs. household as unit of observation, 2) the differences in exposure reference periods across the three waves' questionnaires (past 30 days/recently in Wave 1, past 30 days in Wave 2, past 90-days in Wave 3).

We will note challenges to implementation, including partnership begun before contract award, delays in hiring data collection staff due to lack of clarity on security requirements, delay in Authorization To Operate (ATO), difficulties understanding paid media plan, change in advertising plans prohibiting inclusion of confirmed awareness questions in Wave 2 and mode-specific cueing questions in Wave 2, inability to include selected American Indian Alaskan Native (AIAN) reservations in data collection due to Census Bureau contact protocols, and no web survey in Waves 1 or 2.

We will describe the heavy-up experiment, and note that the results of the experiment are mostly analyzed separately, although some analyses of heavy-up data are presented in this report, where the evaluation objectives could not be satisfactorily tackled with only the main survey data.

E. Complementary Data Sets

This section of the final report will discuss each data set in turn (census paradata, media buy plan, ratings data, data from the Integrated Partner Contact Database (IPCD), data on earned media from the VOCUS database, and absence of administrative data on Census in Schools). Discussions will include any features relevant to the inferences that can be made from incorporating these data into our analyses. Census paradata discussion will cover the process of matching survey households to the Master Address File (MAF).

F. Design Differences from the 2000 Partnership and Marketing Program Evaluation (PMPE)

Key differences from the 2000 PMPE will be discussed, including 1) no Random Digit Dialing (RDD) in Wave 1/2010, 2) better designs for supplemental samples in 2010 3) attempted better handling of paid media in 2010, 4) efforts to separate partnership from paid media (including Heavy-up), and greater mixed mode approach in 2010.

G. Estimation

Since the 2010 CICPE samples are not simple random samples, weights and design-corrected standard errors will be necessary for all the analyses described in this book. Weights will be created for each of the six race/ethnicity groups separately, though the Core sample race/ethnicities (Hispanic, non-Hispanic Black, and non-Hispanic Others) will be more easily combined for analyses than the three supplemental oversamples (Asian, Native Hawaiian and Other Pacific Islanders, and American Indian and Alaska Natives).

The weights will start with a base weight based on selection probability, and adjustment steps based on eligibility, screener response, and interview response. The last weighting step will adjust the weights so that they conform to known population totals. The weights for the Core sample will result in nationally representative estimates. Combining all six race/ethnicity groups will also result in nationally representative estimates, but the variability in the weights will lead to increased standard errors. We plan to explore this issue during our analyses.

All four samples (the Core sample and the three supplemental oversamples) are area-probability samples, which means they are multi-stage stratified cluster samples. Many statistical formulas and packages assume that the data are a simple random sample, so standard errors must be calculated according to the statistical design. Area probability samples usually have larger standard errors than simple random samples, and this ratio is represented by the square root of the design effect, which is defined for variances (squared standard deviations). Design effects for the very similar 2000 PMPE averaged around 2.0, and we expect the design effects for the 2010 CICPE to be similar.

Specialized software is needed to calculate these design-corrected standard errors, but this software is now available for most analyses, and even appears in general-use programs such as SAS. We plan to use the package SUDAAN, which is SAS-compatible, to calculate the standard errors and confidence intervals for the estimates in this analysis report.

Chapter 2.
Defining Sub-populations

A. Hard-to-Count Groups (Race/Ethnicities)

The 2010 CICPE sample design calls for approximately equal sized samples of six hard-to-count groups, defined by race and ethnicity. This section of the 2010 CICPE final report will discuss the design of the core and supplemental samples, as well as provide numbers of cases in each sample type by wave and panel status. Where available, tables will also document the composition of each sample type across detailed categories (e.g., nationalities, tribes, or ethnicities).

B. Audience Segments

The Census Bureau undertook an audience segmentation exercise as part of the planning process for the 2010 Census. Through review of tract-level Census 2000 results, American Community Survey results, and demographic characteristics of households by tracts, the Census Bureau developed an eight-category segmentation of U.S. households. That audience segmentation scheme has informed the planning and implementation of the 2010 CICPE. One table will show the distribution of cases by wave across the eight segments, including panel and non-panel sample members. A second table will show the distribution of cases by sample type across the eight segments. A third table will show by wave whether or not the characteristics of each household would map to the same audience segment as is attributed to its tract. These tables will pertain only to the main survey data (not heavy-up cases). If necessary, a final table will document a collapsed set of audience segments that will be used throughout the report.

We note that households may not share the characteristics that drove their tract's assignment to an audience segment. We will construct a table that indicates for each audience segment the fraction of survey households within that segment that substantially differ from their tracts in propensity to complete the census form. This lays the groundwork for later analyses that hypothesize that households that do not share their tract's characteristics might have weaker response to the communications program than households that do share their tract's characteristics.

C. Heavy-up

The 2010 CICPE final report will not include analyses of the heavy-up experiment per se, but will include some analyses of heavy-up data to supplement key analysis questions of the experiment. Tables showing the numbers of heavy-up cases by sample type and audience segment will be included here.

D. Census Barriers, Attitudes and Motivators Survey (CBAMS)

This section of the report will describe the process of approximately reproducing the CBAMS mindsets using the subset of questions that appear in the 2010 CICPE questionnaire. It will provide mindset distributions by wave for all sample types and all audience segments. For panel cases, it will also provide a transition matrix documenting all individual-level changes (or absence thereof) in mindsets across waves.

E. Clustering by Geography

Use of supplemental data sources such as the IPCD, paid media ratings, and VOCUS tracking of earned media will involve linking data by geography. To provide context for those analyses, we will present some summary statistics on the clustering of cases in geographical areas (such as Designated Market Areas (DMAs), cities, etc.).

F. Other Descriptive Statistics

The survey data include additional characteristics such as: demographic characteristics, media use, community involvement and connectedness, and awareness of government agencies and programs. Brief descriptive statistics will be provided on these characteristics for the main survey sample.

Chapter 3.
Outcomes of Interest

Almost all of the data for measuring our three outcomes of interest will come from the Census Bureau. These include mailback status of each sampled household, paradata on the household relating to mailback participation (date of return, replacement mailings, area of bilingual mailing, etc.), and data from the Non-response Follow-up (NRFU) form and cover sheet (if applicable).

A. Mailback Response Rate

The actual mailback participation status of the survey households will be the primary and most obvious measure of mailback response rate. As needed, a more nuanced measure of latent cooperativeness in the mailback phase may be developed using paradata and survey questions asking about what happened to the form in households that did not mail back (e.g., did anyone open the envelope, did it seem long, etc.). Tables will show mailback rates by wave, by panel status, by sample type, and by audience segment.

B. Reduction of Differential Undercount

Inference on whether the 2010 Integrated Communications Campaign (2010 ICC) reduces the differential undercount will be difficult given the data we will have access to. NORC's proposal provided a road map for how the Census Bureau's coverage data might be used with the 2010 CICPE survey data to discuss the effect of the 2010 ICC on differential undercount. Our understanding is that the coverage files will not be available prior to the end of the 2010 CICPE contract and submission of the final report. The final report will include an updated 'road map' for analyses that could be done once the appropriate data are available. As an interim product, the final report will also consider differences across sample types in responses to the communications program. To the extent that historically less cooperative groups show stronger responses to the program and more cooperative groups show weaker responses to the program, such responses would be interpreted as evidence in support of reduction of differential undercount. The reverse patterns could be interpreted as evidence that the communications program may have exacerbated differential undercount.

C. Cooperation with Enumerators

Cooperation with enumerators is a meaningful construct only for households who have not mailed back the census form prior to the start of NRFU. This outcome will be measured primarily from the NRFU cover sheet and other Census Bureau paradata. Since the Census Bureau expects to enumerate essentially all NRFU households, it is not meaningful to calculate a NRFU response rate for these households. Rather, we will use proxies for 'effort required' to complete these enumerations. For example, we envision using increasing numbers of attempts and proxy completion as indications of poor cooperation with enumerators. We understand that there may be alternative interpretations of these metrics (for example, schedule unavailability could be the reason for large numbers of attempts), and will work with the Census Bureau to identify additional or better measures for assessing households' cooperation with enumerators. The date of final enumeration may also be an indication of cooperativeness (later is worse). Tables showing cooperativeness will show only those cases that were designated for non-response follow-up. These tables will indicate cooperativeness by wave (including panel status), sample type, and audience segment. A single variable will be chosen for this construct for use in later analyses.

Chapter 4.
Exposure to and Experience of the 2010 Integrated Communications Campaign (2010 ICC)

This chapter of the 2010 CICPE final report will begin with a description of our overall strategy for measuring exposure to the communications program. This strategy incorporates: quantitative self-reported data on exposure; qualitative self-reported data on perceptions of and reactions to the communications program; external sources of data such as the IPCD, the Draftfcb media buy plan, and gross ratings points; combinations of these elements to construct component-level exposure measures; and construction of a continuous total exposure measure from these component-level measures.

Tables will document exposure to each program component and the full communications program by sample type, audience segment, wave, and separately for panel members. To explain the construction of these tables, the chapter of the report will review the various elements that contributed to the recommended program exposure measures.

Tables will document paid media/advertising elements reported/recalled least often or most often, and partnership elements reported/recalled least often or most often over time by audience segment and sample type. Discussion will review the difficulties of capturing this information given the communications program design, the less precise information available on partnerships from the IPCD database and on the Census in Schools program (as compared to the paid media measurements), and NORC's inability to include confirmed awareness items in Wave 2.

The survey questionnaires include various items on the qualitative experiences of the program (reactions to ads, whom the respondent spoke with, etc.). Descriptive summary tables will be provided to portray the nature of program exposure.

Tables incorporating supplemental data sources will provide measures of communications program exposure as alternatives to the self-reported measures.

In addition to total exposure calculations, we will document the correlation of exposure across components among individuals. This correlation will affect the extent to which it is possible to separate out component-level contributions to outcomes of interest.

Two methodological investigations will also be reported here. The first will focus on false ads included in Waves 1-3 in order to improve the accuracy of our estimates of paid media exposure. The second will study confirmed awareness items in Wave 3 and the extent to which estimates of paid media exposure might differ based on the inclusion or exclusion of these items.

Summary statistics on heavy-up exposure will also be reported here, primarily to see the differences in correlation across component-level exposure.

Chapter 5.
Knowledge, Attitudes and Beliefs about the Census

Knowledge, attitudes and beliefs are a key focus of the structural model for the 2010 CICPE. Two particular analytic objectives are to 1) understand the extent to which knowledge, attitudes and beliefs about the census are related to intent to participate and/or actual census participation; and, 2) investigate the relationship between individuals' exposure to various 2010 ICC components and their (changes in) knowledge, attitudes and beliefs about the census. A third critical topic is the relationship between intent to participate and actual census participation, although that is less an evaluation topic than information of relevance to the Census Bureau's understanding of the mechanisms that lead to census participation and their relevance to future planning efforts. This last topic will be discussed in Chapter 9 of the final report. Specific evaluation objectives addressed by this chapter include objectives 4, 5, 6 and 9 (see Section 1c).

A. Overview of Data Items

Several of these items, especially regarding awareness and intent to participate, will have been presented briefly in Chapter 2 in reconstructing the CBAMS mindsets.

In general, weighted frequencies will be presented at the item-level by wave, sample type, audience segment, and among panel only for each set of variables (awareness, positive attitudes, negative attitudes, etc.). These tables will include summary variables (scales or constructed measures) for each set. Remaining analyses will typically use the summary measures. For example, the knowledge items may be summarized as # of true items reported correctly and # of false items reported correctly.

Knowledge, attitudes and beliefs, including awareness of the census and intent to participate, will primarily be drawn from the survey data.

Awareness

4C. Have you ever heard of the Census?

5. The Census is the count of all the people who live in the United States. Have you ever heard of that before?

13. In general, how familiar are you with the way Census data impacts you and your community. Would you say…very familiar, etc.?

Intent

9. If the Census were held today, how likely would you be to participate? By participate, we mean fill out and mail in a Census form. Would you say you… definitely would, etc.?

Knowledge (true or false)

14. So far as you know, does the law require you to answer the census questions?

15. People have different ideas about what the Census is used for. I am going to read some of them to you. As I read each one, please tell me by indicating yes or no whether you think that the Census is used for that purpose. Is the Census used…

    a. to decide how much money communities will get from the government?
    b. to decide how many representatives each state will have in Congress?
    c. to count both citizens and non-citizens?
    d. to determine property taxes? – *false*
    e. to help the police and FBI keep track of people who break the law? – *false*
    f. to help businesses and governments plan for the future?
    g. to locate people living in the country illegally? -- *false*

Attitudes and Beliefs

12. Overall, how would you describe your general feelings about the Census? Do you feel... highly favorable, etc.? *Qualified also by 12a (certainty about favorability).*
16. Next, I'm going to read some opinions about the Census. As I read each one, tell me if you strongly agree, agree, disagree, or strongly disagree with each of the statements:
    16A. Filling out the Census will let the government know what my community needs.
    16B. The Census is an invasion of privacy. – *negative*
    16C. The Census Bureau's promise of confidentiality can be trusted.
    16D. I am concerned that the information I provide will be misused. – *negative*
    16E. Taking part in the Census shows I am proud of who I am.
    16F. My answers to the Census could be used against me. – *negative*
    16G. Answering and sending back the Census matters for my family and community.
    16H. The government already has my personal information, like my tax returns, so I don't need to fill out a Census form. – *negative*
    16I. I just don't see that it matters much if I personally fill out the Census form or not. – *negative*
    16J. It takes too long to fill out the Census information, I don't have time. - *negative*

B. Associating Program Exposure with Knowledge, Attitudes and Beliefs

This section of the report will show cross-tabulations of high vs. low exposure by program component and total exposure against these summary measures (awareness, intent, knowledge of true items, knowledge of false items, as well as positive attitudes and negative attitudes). We will also tabulate *changes* in awareness, intent, knowledge of true items, knowledge of false items, positive attitudes and negative attitudes for the whole sample across waves. We will then show changes in these measures for only the panel cases, again segmenting them by their levels of program exposure.

C. Changes from the 2000 PMPE to 2010 CICPE

The subset of items appearing in the 2000 PMPE as well will be used to report on comparisons in these measures between the two censuses.

D. Regression Analyses Linking ICP Exposure with Changes in Knowledge, Attitudes and Beliefs

Finally, we will report some multivariate regression results assessing the extent to which exposure to different program components were related to changes in knowledge, attitudes and beliefs. These analyses may also be repeated on the heavy-up sample.

Chapter 6.
Associating 2010 ICC Exposure with Census Participation

This chapter of the 2010 CICPE final report will document associations between communications program exposure and census participation, primarily through cross tabulations. Associations will be examined by program component (paid media, partnership, earned media, census in schools) and for mailback response rate and cooperation with enumerators. Where possible, tables will be interpreted for implications on differential undercount. Tables will be presented by audience segment, wave, sample type, and panel status.

We note that the analyses proposed here are simply associations by program component, which will be straightforward to construct, if difficult to interpret. As discussed in Chapter 1, the integrated nature of the 2010 ICC makes it very difficult to cleanly separate effects of the different components of the campaign (partnership, paid media, earned media, etc.). That difficulty will limit the interpretation of associations in this chapter.

The chapter will also include a discussion of advertisements, programs, and events (including breaking news events) outside of the 2010 ICC that may have had an effect on respondent attitudes and behaviors. The analyses relevant for this discussion will be developed once other analyses have been completed and more is known about highly salient events that may occur during the first four months of 2010.

Chapter 7.
Multivariate Analytical Models

Previous analyses in the report will focus on individual questions in the survey and address specific issues. In this part of the final report we will present more comprehensive analyses making simultaneous use of a number of the survey variables. The goal is to make more overarching statements about how the communications program affected people and influenced their participation.

The analyses will be predicated on the following strategy. The structural model described in Section 1b will provide an overall blueprint for the analyses. However, we will approach this model sequentially rather than as "the model" to be tested. Specifically, we approach the analyses on the assumption that there is no one model that can characterize the communications program in its entirety. Rather we need to consider different scenarios for how the communications program could have affected people and to build a more selective model based on these analyses. Moreover, it is likely that the communications program affected people in different ways, so that the best approach is also to incorporate individual differences into the model.

We will thus begin by considering alternative scenarios for characterizing the impact of the communications program. Our preliminary framework envisages five different scenarios.

Minimal impact scenario. Here we will consider the possibility that the communications program had little or no impact in actually influencing key census behavior dependent variables. In effect, under this scenario, people already were either likely or unlikely to participate based on background variables independent of the program. We would develop a model for this scenario based on variables having to do with background knowledge about government programs, general social and political participation, patterns of general media use, and education. Variables such as advertising exposure reported would be viewed as consequences of these variables.

Attention/familiarity scenario. Under this scenario, the communications program may have raised awareness of the census and created a sense of familiarity that affected behavior without any deeper cognitive impact. In effect people may have processed the fact of the 2010 Integrated Communications Campaign (2010 ICC) rather than the actual messages being delivered. Key variables would include measures of exposure to the ads and false recall, with more cognitive variables used as covariate controls.

Belief acceptance/change scenario. Under this scenario, people would have responded to the advertising messages in ways consistent with their underlying beliefs and/or changed their beliefs.

Social influence scenario. It is possible that people responded primarily to the partnership activities including the schools program and that responses to the media campaign merely reflect

this. Relevant variables consist of partnership exposure and the overall level of social interaction relevant to the census.

Earned media scenario. Exposure to media coverage may have been the key determinant of census behavior. This coverage may have produced a sense of currency and credibility/controversy that affected people, with campaign and partnership reactions reflecting this.

A Three-Stage Analysis Strategy

For each scenario we will explore analytical models and compare the results across the above scenarios. Within each scenario, we will try to fit the best model we can. We will then try to assess the goodness of fit and explanatory power of the model, and we will compare the alternative models/scenarios. For instance, with the belief acceptance scenario, we would examine predictions about how census-related beliefs are related to message reactions as might be expected given the scenario. The panel part of the data will also be used for model exploration in this same vein. This constitutes the first stage of the analysis.

In the second stage we hope to build a case (for each scenario) for including variables in a comprehensive second stage model that allows for multiple impacts. Such a model necessarily faces problems of causal interpretation and multi-collinearity, so that the strategy of first building and evaluating more limited models should be helpful in selecting the variables to be included.

A final, third stage of the analysis will be based on the assumption that different people will have been affected by the communications program in different ways. A comprehensive model may capture aggregate effects but it will also be informative to examine individual differences. The approach we will take is to use each of the stage-one scenario models to characterize individual cases (respondents). For instance, a respondent's census behavior may reflect more of the social influence and attention/familiarity impact than the belief acceptance.

We will construct second order variables reflecting such individual differences in response each respondent receiving a score based on each scenario. These can be compared to each other to provide further insight into the relative impact of the different scenarios. They can also be used to classify respondents and thus measure the number of people impacted in different ways. These classifications can be profiled on relevant descriptive variables (e.g. demographics) and compared to campaign relevant segments (e.g. hard to reach or CBAMS membership types).

It is anticipated that these analyses will be cumulative in the sense that they will lead to comprehensive but nuanced general conclusions about the communications program.

Chapter 8.
Return on Investment

Our overall approach to measuring return on investment in the 2010 Integrated Communications Campaign (2010 ICC) is to associate program investment with census outcomes. The final step of linking those census outcomes with dollar values, we would leave to the Census Bureau to complete, perhaps closer to the time of the 2020 Census.

Most of the steps of this approach will have been completed as discussed in previous chapters, especially the multivariate regressions in Chapter 7. For example, we will have:

(1) estimated the effect of changes in awareness on mail return rate

(2) estimated the effect of changes in awareness on average number of interviewer visits per non-responding household (or some alternative measure of cooperation with enumerators)

(3) estimated the effect of changes in Gross Rating Points (GRP) on awareness

(4) estimated the effect of changes in GRP on mail return rate

(5) estimated the effect of changes in GRP on the average number of interviewer visits per non-responding household.

Gross Rating Points (GRPs) are a conventional unit used by advertising researchers to measure a population's opportunity for exposure to media content (Farris & Parry, 1991). GRPs are an estimate of underlying reach and frequency of a media execution such as a television program or advertisement running during a program. GRP estimates reflect the percentage of a given population (e.g., in the DMA of Los Angeles) and frequency with which that population had the opportunity to be exposed to a program or advertisement.

The primary task in this chapter will be to depict several scenarios to assist in interpretation of the analyses above in a return on investment context.

We will first bring in information on the Census Bureau's investment in various components of the communications program. (We have requested such information; our ability to complete this step of the analysis will be determined in large part by the quality and nature of the information available on total Census Bureau investments in the 2010 ICC. It has been fairly noted that a true calculation of investments would represent only the incremental investments above and beyond what would have been paid for activities in the absence of an Integrated Communications Campaign)

To illustrate return on investment, we will take the determined gains or losses (e.g., in the mail return rate) due to 2010 ICC components, determine the amount of money saved, and compare

these savings against the costs incurred by the 2010 ICC component. We would provide such illustrative calculations by sample type, and for mail return rate and cooperation with enumerators. As an example for the media buys, we will try to construct an analysis to predict the marginal improvement in response rate from a marginal increase in GRP. We will not have the data to construct a full curve, but rather we will try to approximate the slope of the curve around the point of actual GRPs achieved. Under alternative cost scenarios, where we vary how much money the marginal increase in GRP costs and how much money a marginal increase in response rates saves, we can calculate the net savings (or lack thereof) from the purchase of the GRPs. Using both self-reported and supplemental source data (e.g., GRPs or metrics from the IPCD) offers opportunities for conducting sensitivity analyses of these results.

The same calculations may be provided based on the heavy-up sample estimates as well.

Chapter 9.
Other Methodological Issues

A. Panel Composition

For Waves 2 and 3, the samples will be a mix of panel cases and non-panel cases. The panel cases will be selected from the Wave 1 completes so that the panel is representative of the Wave 1 completes. Of course, Wave 1 has non-response and this will carry over into Waves 2 and 3. This section of the 2010 CICPE final report will analyze the panel samples by race/ethnicity and possibly other variables in terms of response rates and other characteristics.

B. Conditioning Effect of Panel Participation

Panel conditioning is a major measurement error unique to panels. It refers to the phenomenon where panel participation in repeated interviews changes respondents' behavior and attitudes *or* their report of their behavior and attitudes. For this evaluation, some of the Wave 1 respondents will be invited back to participate in the second and third wave. Their participation in the first wave interviews informs them about the survey topic, sponsor, tasks and burdens associated with the survey and what to expect with the next interviews. All these factors are expected to affect how they behave towards the next interviews. We will examine the ways in which panel respondents who have participated in more than one wave behave differently from those who are new to each wave. Specifically, we are interested in examining how the experience of a prior interview will affect respondents' decision to participate in later waves, reporting of knowledge, attitudes, and beliefs about the census, reports of exposures to various program components, intention to return the census form, and the actual decision to return the census form.

To measure panel conditioning, we can compare Wave 2 fresh cases with Wave 2 panel cases with the same personal characteristics, and Wave 3 fresh cases with Wave 3 panel cases with the same personal characteristics. In addition, we can compare Wave 3 panel cases that did not complete the Wave 2 interview with Wave 3 panel cases that did complete the Wave 2 interview. We can also compare panel cases that completed all 3 waves with panel cases that completed on only two waves (either Wave 2 or Wave 3). Some variables that might be most likely to show conditioning effects are: communications program exposure questions (Wave 1 items Q17a-Q17f, Q18, Q20a-Q20f, Q21a, Q21b, Q22g, Q24a-f, Q25a, Q26b), knowledge questions (Q15a-Q15g), attitudes/beliefs questions (Q16), intention to return census form (Q9, Q10). In addition, questions involved in skip patterns tend to be affected by panel conditioning as well such as Q8, Q22, and Q22d. The actual mail return status will also be compared between panel respondents and fresh cases.

C. Demographic or Behavioral Differences by Mode

The sample will be mixed mode, with some telephone interviews (TI) and some personal interviews (PI). Still others will be self-administered questionnaires (SAQ) by paper-and-pencil in Wave 2, and potentially both paper-and-pencil and web in Wave 3. The administrations by telephone and SAQ are cost-savings measures and also provide respondents additional convenient means of survey participation. The sample as selected is a sample of addresses, and those addresses with known telephone numbers will be included in the TI part of the survey. Nonparticipants in the TI survey and addresses without known phone numbers will be included in the PI part of the survey. Panel participants in Waves 2 and 3 will first be approached for SAQ completion, then followed-up by phone or in-person as needed. An open question is whether there are mode effects in the survey. Differences in statistics based on the TI, PI and SAQ data could reflect mode differences or they could reflect differences in the subpopulations receiving the various modes. Similarly, lack of differences in statistics based on the data by mode could reflect offsetting mode differences and differences in the subpopulations receiving the various modes.

D. Self-reported vs. Actual Participation

Respondents to the survey will report whether their household's census questionnaire was returned by mail or not. These reports are subject to two major kinds of errors, which go in opposite directions. Some respondents whose questionnaire was not returned by mail will report that it was, in order to report a behavior (civic participation) that is seen as positive in most social contexts. On the other hand, some respondents may be unaware whether or not their household's questionnaire was returned by mail (perhaps by another member of their household or because they did not recognize the completed document as the subject of our survey questions), and will report that an inaccurate mailback status inadvertently. We will compare self-reported versus actual participation, and if the discrepancies are significant, investigate demographic and other correlates of the discrepancy. In addition, we will use multivariate regression techniques to measure the relationship between reported intent and the two participation measures (self-reported and actual) and describe the natures of those two relationships. One would hypothesize that intent might better predict self-reported participation than actual participation.

E. Non-response Bias

Our survey will be subject to non-response, and it is reasonable to anticipate that non-respondents to the survey will differ in their knowledge of, attitudes to, and behavior in the decennial census. Comparisons between respondent households and non-respondent households can be carried out on census enumeration data and paradata, provided that the matches are done (as noted in 9h) and the data are provided. Examples of paradata would include whether a mail return was received, how many reminders were sent, and how many visits by an interviewer took place. Using the enumeration data and paradata, we can compare the non-respondents and

respondents to our survey. We can also use those data to develop propensity models and other approaches to weighting adjustments for non-response. This exercise will be more straightforward to complete on the Core Survey sample (where all households are eligible for survey participation) than on the supplemental samples (where the vast majority of households are ineligible for survey participation by virtue of their race not being the targeted race of the sample component). At least in the Core Survey sample, all sampled addresses' census enumeration data and paradata will be analyzed for non-response bias, whether or not they were successfully screened or interviews for the 2010 CICPE.

One possible approach will be to assign different non-response propensity scores to our own survey data (for example, based on date of completion, number of attempts required, or other paradata variables). To the extent that key analytic results are stable across non-response propensity, we would surmise that the effect of total non-response on the 2010 CICPE results may be more modest. To the extent the results differ by non-response propensity even among survey respondents, we might infer that non-response to the 2010 CICPE is a greater factor in the generalizability of evaluation results to the rest of the population.

F. Self-reported vs. External Data on Exposure

There are two basic types of advertising exposure measures: 1) self-report and 2) environmental measures. Self-report measures can be divided into 1) recall and 2) recognition measures. Recall measures capture individuals' aided (with verbal or similar prompts) or unaided (without prompts) awareness. They rely on cues to memory such as a brief description of what an advertisement was about or some feature of it. To demonstrate exposure, respondents then provide a more detailed self-report (correctly describing some detailed feature of the ad) to demonstrate recall. Recognition measures rely on actual depiction of the advertisements through presentation via media such as TV, Internet, radio, or print. Respondents view/hear/read the ads and indicate their recognition of the ad. Frequency of exposure may be asked for both recall and recognition as well.

Environmental measures use extant data about the delivery of an advertised message, typically through mass media, in a given geographic area (e.g., a Designated Market Area(DMA)).

It is important to distinguish between *recognition* and *recall* of 1) messages and 2) advertising executions. A myriad of recent studies provide evidence that advertising campaigns can be effective in changing health-related attitudes, intentions, and behavior such as tobacco use (Farrelly et al., 2005; Hersey, Niederdeppe, Evans, et al. 2005; Hersey, Niederdeppe, Ng, et al. 2005; Sly, Trapido, and Ray, 2002) and other health or problem behaviors (Hornik, 2002). Studies have examined measures of exposure to and/or memory of televised advertisements. Measures of *recognition* typically used in laboratory forced exposure designs have participants view an ad and indicate whether they have seen the message before, whereas *recall* measures provide study participants with only a minimal verbal cue (Evans et al, 2009; Singh and Rothschild, 1983). These studies have shown that recognition and recall measures are strongly correlated, and researchers generally agree that both measures require respondents to access their long-term memory for at least minimal remnants of ad exposure (Southwell et al., 2002).

In recent years, recognition measures have also been used in field studies in which respondents are shown ads that have recently (e.g., in the past 1-2 months) aired on TV, on radio, or appeared in print in their DMA. They are then asked to indicate whether they have seen the ad(s) and if so with what frequency on an ordinal scale.

For the 2010 CICPE, we will use both measures of recognition and GRPs for paid advertising in the media campaign. For advertising, this includes both aided and unaided awareness. Adapting from media measures, we have created aided awareness measures for new media, partnership, and word-of-mouth. Little is known about validity and reliability of these measures from previous research. We also have access to external data sources that may provide varying levels of valid and reliable estimates of exposure sources. Table 9 provides a summary of each domain of measures.

Table 9. Summary of exposure measures for 2010 CICPE

Source of exposure	Recall	Recognition	Environmental estimate
Advertising	General campaign aided awareness (not of specific ads) False recall of specific ad (unaided)	Wave 3 only – Show recently aired ads Ask for recognition Ask for frequency of exposure	Gross rating points (GRP).
New media (e.g., online social networking)	General campaign awareness. Reported frequency of exposure by source.	N/A	Web metrics (hits, etc.)
Partnership events	General campaign awareness. Reported frequency of exposure by source.	N/A	Census partnership database.
Census-in-schools	General campaign awareness. Reported frequency of exposure by type of communication	N/A	Presence of children in K-12 school.
"Word-of-mouth"	Reported participation frequency.	N/A	N/A

Thorny issues arise for less well studied, harder-to-quantify sources of exposure including partnerships, new media, Census-in-schools, and "word-of-mouth." There are two major exposure issues to address in the evaluation: 1) lack of complete overlap in types of measures across exposure sources; and 2) lack of knowledge of validity and reliability of measures outside of advertising recall domain. Partner organizations have been provided material that looks similar to the paid media materials, partners have shown Census ads to their membership and constituencies, and have privately sponsored advertisements of their own. Using the IPCD, the NORC evaluation team will attempt to document potential unique sources of exposure from partnership activities and develop a partnership exposure index score as an estimate of the

independent exposure of respondents to partnership activities by DMA (similar in format to a GRP estimate). It should be recognized that this procedure has not been previously validated, and will need to be developed and tested for possible use in the ICP evaluation.

Finally, we will compare self-reported 2010 CICPE campaign exposure obtained from recognition measures to environmental exposure obtained from GRP estimates. We will conduct two sets of multivariable regression models, controlling for various geographical and demographic characteristics of survey respondents, using in one case the recognition and in the other GRP measures as independent variables. We will compare the observed regression coefficients for outcomes of interest (e.g., changes in beliefs about the Census, intent to mail back the form, and actual mail back behavior) and determine whether there are statistically significant differences in outcomes based on use of recognition versus GRP measures. We will conduct this methodological study for each component of the campaign for which sufficient data are available to do so.

G. Mode-specific Cueing (Wave 3 only)

To measure mode-specific response to Census advertising, we will follow validated techniques to measure encoded advertising exposure (Southwell et al, 2002). Exposure will be measured by testing for respondent recognition of Census ads that have recently run on major media outlets, including those targeting specific racial/ethnic groups. To assess recognition, respondents to the web-based and paper-and-pencil survey instruments will first be shown still images taken from the most frequently aired television advertisements. Both targeted and general population ads will be excerpted. This procedure has been used in several recent marketing campaign evaluations, and we will model these procedures in this study (Evans et al, 2009; Davis et al, 2009).

For the telephone interviews, respondents will hear radio ad clips. Respondents who were unable to hear the radio ad will instead be provided a text script of the audio from that ad. Once participants viewed and listened to each ad, they will then be asked how frequently they had seen/heard/read each ad on the specific medium used in that survey, based on a scale ranging from ‘very often’ to ‘never’. Respondents who indicated having seen the ad at least ‘rarely’ will be considered to have recall of that ad. This procedure has been validated as a method of measuring encoded ad exposure (Southwell et al., 2002).

After viewing/hearing/reading each ad, respondents will immediately be asked a series of questions to assess specific reactions to the ad. These measures include the following:

- Was the ad convincing?
- Did the ad grab your attention?
- Did the ad give good reasons to mail back the Census form?
- Did you talk to friends, family, or other people about the ad?
- A series of questions about what the ad said to the individual (take away messages internalized by the respondent)

H. Match Rates to Master Address File (MAF), Differences Across Waves

Many of our most important analyses use actual Census participation data to measure the impact of the 2010 ICP. However, we will have to match our interview addresses to the MAF in order to obtain this Census participation data. We expect this matching operation to experience difficulties, especially in rural areas. This section will examine the match rates by sample type, wave, and other variables.

I. Language of Interview

The 2010 CICPE interviewing staff is collecting interviews in English, Spanish, Russian, Korean, Vietnamese and Chinese (simplified). While English and Spanish interviews can be completed by CATI or using translated paper and pencil instruments, the other four languages are available only in translated paper and pencil forms. In Wave 1, we have fewer interviews in languages other than English and Spanish than we expected. This section will show the number of interviews in each language and will analyze and compare the interviews in languages with enough interviews to analyze. At a minimum, the English, Spanish, and Other Language interviews will be compared.

J. Incentive Experiment Results

In order to better understand the impact of incentives on refusals, we are carrying out an incentive experiment on Computer Assisted Telephone Interview (CATI) refusals. Cases that were determined to be eligible for the study but which continued to be refusals were considered for the incentive experiment. In Wave 1, incentive-eligible CATI (pending) refusals received an advance letter or learned from the telephone interviewer that they could receive $20 cash payment on interview completion (previously, no incentive was offered). After eliminating out of scope/ineligible, hostile refusal, and 'Take Me Off the List' cases, there were 640 cases eligible for the incentive. In order to maximize the benefit of the incentive while still carrying out the experiment, we selected 64 cases to not receive the incentive, while the other 576 cases received the incentive. This section will analyze the response rates for the two groups and see how successful the incentive was in increasing the response rate.

K. Late Wave 1 Cases

Wave 1 data collection continued through January 16, 2010, although advanced press about the paid media campaign built quickly starting with multiple appearances by the Director of the Census Bureau on morning news shows on Monday, January 4, 2010, and including detailed coverage of the paid media campaign following a press briefing on January 14, 2010. Wave 1 interviews completed on or after January 4 will be reviewed for sharply increased reports of paid media and earned media exposure relative to earlier Wave 1 interviews. If such increases are found, then cases completed after a certain date will be excluded from selected 'baseline' analyses of Wave 1 data.

References

Davis, K.C., Farrelly, M.C., Messeri, P., & Duke, J. (2009).The impact of national smoking prevention campaigns on tobacco-related beliefs, intentions to smoke and smoking initiation: Results from a longitudinal survey of youth in the United States. *International Journal of Environmental Research and Public Health., 6*, 722-740.

Evans, D. W., Uhrig, J., Davis, K., & McCormack, L. (2009). Efficacy methods to evaluate health communication and marketing campaigns. *Journal of Health Communication, 14,* 314-330.

Farrelly, M. C., Davis, K. C., Haviland, M. L. Messeri, P., & Healton, C. G. (2005). Evidence of a dose-response relationship between "truth" antismoking ads and youth smoking prevalence. *American Journal of Public* Health, *95*, 425-432.

Farris, P.W., and Parry, M.E. (1991). Clarifying some ambiguities regarding GRP and average frequency. *Journal of Advertising Research, 31*, 75-77.

Hersey, J.C., Niederdeppe, J., Evans, D. W., Nonnemaker, J., Blahut, S., Holden, D., Messeri, P., & Haviland, M. L. (2005). Empirical articles – The theory of "truth": How counterindustry media campaigns affect smoking behavior among teens. *Health Psychology, 24,* 22-32.

Hersey, J. C., Niederdeppe, J., Ng, S. W., Mowery, P., Farelly, M., & Messeri, P. (2005). How state counter-industry campaigns help prime perceptions of tobacco industry practices to promote reductions in youth smoking. *Tobacco Control, 14*, 377-384.

Hornik, R. C. (2002). *Public health communication: evidence for behavior change.* Mahwah, NJ: Erlbaum Associates.

Sly, D., Trapido, E., & Ray, S. (2002). Evidence of the dose effects of an antitobacco counteradvertising campaign. *Preventive Medicine, 35,* 511-518.

Southwell, B., Barmada, C. H., Hornik, R., & Maklan, D. (2002). Can we measure encoded exposure? Validation evidence from a national campaign. *Journal of Health Communication, 7,* 445-453.