UNITED STATES DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. Census Bureau
Washington, DC 20233-0001

April 30, 2012

2010 CENSUS PLANNING MEMORANDA SERIES

No. 191

MEMORANDUM FOR The Distribution List

From: Arnold Jackson *[signed]*
Acting Chief, Decennial Management Division

Subject: 2010 Census Paid Advertising Heavy-Up Experiment Evaluation Report

Attached is the 2010 Census Paid Advertising Heavy-Up Experiment Evaluation Report. The Quality Process for the 2010 Census Test Evaluations, Experiments, and Assessments was applied to the methodology development and review process. The report is sound and appropriate for completeness and accuracy.

If you have questions about this report, please contact Nancy Bates at (301) 763-5248.

Attachment

April 24, 2011

The Paid Advertising Heavy-Up Experiment

U.S. Census Bureau standards and quality process procedures were applied throughout the creation of this report.

Nancy Bates
Kristin McCue
Michael Lotti
Research and Methodology Directorate,
Center for Economic Studies,
Accretive Insights

United States Census Bureau

United States Census 2010

This page intentionally left blank.

Table of Contents

Section **Page**

List of Tables

Table Name **Page**

List of Figures

Figure Name **Page**

Executive Summary

Background

To evaluate the level of investment in the paid advertising component of the 2010 Census Integrated Communication Campaign, we conducted an experiment in eight matched-pair geographic areas known as designated market areas. These pairs were matched on criteria such as race and ethnic composition, hard-to-count scores, presence of media outlets, and mail return behavior from Census 2000 and the American Community Survey. On average, the areas represented population sizes with adults aged 18 and older of around 500,000. We randomized selection of one area from each pair to receive a "heavy-up" dosage of paid television and radio advertising during the awareness and motivation phases of the campaign. The goal was to increase the treatment area media spending by 100 percent compared to the control areas.

We used two primary data sources to evaluate the experiment: a pre and post-census sample survey and operational Census data from the sixteen designated market areas. The sample survey interviews were part of another 2010 Census evaluation – the Census Integrated Communication Campaign Evaluation. The sample for our evaluation consisted of 1,924 pre-Census and 2,046 post-Census interviews located within the sixteen designated market areas (referred to as wave 1 and wave 3 interviews). The interviews measured levels of self-reported exposure to the Integrated Communication Campaign as well as knowledge, attitudes, and opinions about the Census. Addresses from the sample interviews were matched to Census operational data to create an additional variable indicating whether the household completed and mailed back a census form.

The second data source, the Census operational extract, included all households located within the counties that comprise the sixteen designated market areas. It contained information such as Census mail return status, cooperation during Nonresponse Followup, whether a household was designated to receive a replacement form, and the 2010 Census data collected for the household (for example, race of householder, age of householder, and household size).

Prior to addressing the major research questions of interest, we sought to understand how successful we were in increasing the paid advertising in the treatment areas. The heavy-up experiment was limited to increased advertising as part of the Diverse Mass audience plan and relied on local (as opposed to national) media buys to deliver the extra dosages. A media buy is a purchase of media with specific, controlled requirements for schedule through the day and over weeks, audience characteristics, media characteristics, and program genre characteristics. This placed restrictions on the experiment. Analysis of post-Census expenditure and gross rating point audit data indicated the experiment mostly fell short of the goal of increasing paid advertising by 100 percent in the treatment areas. The audit data also revealed that, even in the *control* sites, the combined total audience paid advertising reach and frequency was very high in some sites – a result of advertising from the targeted Black audience plan, the national Hispanic audience plan, and last minute media buys as part of the Rapid Response Program. These two factors – potential saturation and inability to double the advertising dosages – are limitations that likely affected our results.

Results

How did changes in level of paid media investment relate to changes in Census awareness, knowledge and attitudes and Census advertising recall?

- The survey indicated large and significant increases in self-reports of *any* paid ad exposure before versus after the 2010 Census Integrated Communication Campaign, however, the increase in the treatment areas was not significantly different from the increase in the control.
- Likewise, the *frequency* of self-reported paid ad exposure increased significantly between waves but the increase in the treatment was no different from the increase in the control.
- Self-reported campaign exposure to various media *types* (for example, television, radio, newspapers, and internet) varied little between treatment and control areas. One exception was radio where a larger increase among the treatment group was marginally significant.
- The increase in self-reported exposure to *both* paid media and earned media was significant between waves, but again, the increase was not different between the treatment and control sites.
- The treatments sites showed a significant increase in both ad recall and recall of relevant details of the paid advertisement titled "Frank" which was singled out for detailed investigation. Respondents in the treatments sites also reported a higher frequency of seeing "Frank".
- Of four attitude indices constructed (Privacy and Confidentiality Concerns; Census is Necessary/Important; Census Used for Tracking/Policing; and Census Used to Allocate Resources) only one changed significantly between waves – the index mean for Census is Necessary/Important became higher. However, the mean difference between treatment and control for this index was not significant.
- Taken together, we conclude the extra investment of advertising did increase content recall of the messages, but did not result in higher outcomes, such as awareness, knowledge, or positive attitudes about the census. One explanation is that the base level of paid media produced large increases in awareness, knowledge and positive attitudes making any effect of the heavy-up dosage difficult to detect.

How did changes in level of paid media investment and activity relate to changes in the mail return rates?

- Using census operational data for all households located within the designated market areas in the experiment, we found there was no overall practical difference between the treatment and control site mail return rates prior to nonresponse followup. The combined mail return rate from households located in mailout/mailback areas across all control sites was 75.3 percent compared to 74.9 percent in the treatment sites.
- Within pairs, we found that three pairs had higher mail return rates in the treatment sites, three were higher in the control sites, and two had negligible differences between them.
- Using multivariate modeling to account for potentially confounding differences across designated market areas within a pair does not change the conclusion that there does not

appear to be any significant difference in mail return rates associated with the heavy-up treatment.

- Separate model estimates for non-Hispanic Whites, non-Hispanic Blacks, and Hispanics each show no significant treatment effect. When we estimate the models separately by audience segment, we find marginally significant positive effects of increased advertising for the Single Mobile group, but not for any of the other segments that are represented in the experimental designated market areas.
- Overall, the percent of *survey* respondents who completed and mailed back a census form prior to nonresponse followup was not significantly different between the treatment and control sites (66.7 percent versus 64.0 percent, respectively). Likewise, the difference in mail return rates between treatment and control sites was not significant for any of the race/ethnic groups we examined (White, non-Hispanic; Black, non-Hispanic; and Hispanic, any race).
- We did find that mail return rates among *survey* respondents were significantly higher in the treatments sites for two of the 2010 Census Integrated Communication Campaign audience segments – the Advantaged Homeowner and Economically Disadvantaged segments.

Do experimental findings imply an ultimate savings in field nonresponse followup costs as a result of doubling the up-front investment in paid media?

- In this test, there was no significant increase in the mail return rate for the treatment sites. Therefore, it is not possible to infer a saving in nonresponse followup costs.

Recommendations

- Continue in-market testing of variables subject to Census Bureau control in the communication strategy, but do not repeat a heavy-up only design.
- Design factorial experiments that vary levels of partnership, earned media, and paid media.
- Design tests to include other variables in the paid media mix such as social media and online advertising.
- Design tests that use moderately sized designated market areas not receiving local targeted advertising plans.
- Design measures to more precisely understand impact of the partnerships and level and valence of earned media in the test locations.
- Include an in-market tracking survey as part of the design.
- Use measures for advertising exposure and frequency that do not depend upon self-reports.

This page intentionally left blank.

1. Introduction

1.1 Purpose of Study

As part of the 2010 Census Integrated Communications Campaign (ICC), the Census Bureau contracted with an advertising agency to develop and deliver a multi-million dollar paid advertising campaign. One of the main goals of the campaign was to increase public awareness about the 2010 Census and increase the number of households who participated by completing and returning a census form by mail.

The primary purpose of the paid advertising heavy-up experiment (PAHUE) is to assess the impact that paid advertising had on mail return rates (MRR) during the 2010 Census. For this experiment, the primary research questions are:

1. How do changes in level of paid media investment relate to changes in Census awareness, knowledge and attitudes and Census advertising recall?
2. How do changes in level of paid media investment and activity relate to changes in the mail return rates?
3. Do experimental findings imply an ultimate savings in field Non-Response Followup costs as a result of doubling the up-front investment in paid media?

This experiment will complement a larger evaluation of the 2010 ICC being conducted under contract to the National Opinion Research Center (NORC).

1.2 Background

Paid advertising was used for the first time in a U.S. Decennial Census in 2000. An independent evaluation was performed by the NORC. This evaluation used self-reported measures of ICC exposure to predict propensity to return a census form. However, in Census 2000 there were no controlled experiments carried out to vary the planned "dosage" of paid media administered in one geographic area versus another. Consequently, the NORC evaluation could only analyze variations in self-reported exposure at the household-level as it related to variations in mailback behavior at the household level. The NORC evaluation concluded that the ICC was effective in increasing Census awareness and positively changing attitudes and beliefs about the Census. The evaluation had a harder time making a causal connection between reported ICC exposure and an increased likelihood of actually completing and mailing back a census form (Wolter et al, 2002).

It has been previously recommended that the Census Bureau undertake an in-market controlled test during the 2010 Census to assess the impact of paid media on MRR. This recommendation was made in several Census 2000 evaluation reports, written recommendations from the Census Advisory Committees of Professional Associations, and in recommendations from an independent Academic Assessment Panel (see U.S. Census Bureau, 2009; Academic Assessment Panel, 2009; National Research Council 2004; Edwards and Wilson, 2004; Wolter et. al, 2002).

To act on these recommendations, the Census Bureau conducted a controlled experiment in the 2010 Census. Specifically, eight matched-pair sites were selected and one site from each pair

was randomly selected to receive an increase in the amount of paid advertising "dosage" while the other half received the normal "dosage" applied to areas not included in the experiment.

2. Methodology

2.1 Site Selection

Advertising response tests use pairs or groups of markets that are matched on demographic and behavioral characteristics. The matching was done using the geographic unit used to purchase broadcast media -- geographic areas known as designated market areas (DMAs). Nielsen Media Research defines 210 DMAs for use in TV media measurement. DMAs roughly correspond to metropolitan areas, but also include those surrounding counties where the largest share of viewing is to stations located in the metropolitan area. This experiment matched DMAs in terms of:

- Population size
- Composition of behavioral cluster segmentation used for the ICC (Advantaged Home Owners, All-Around Average, Economically Disadvantaged, Ethnic Enclave, and Single/Mobile[1])
- Factors that make up the hard-to-count (HTC) scores
- Racial and ethnic composition
- Presence and access to media outlets including availability of local targeted cable, print, and original programming
- A level of media spill-in below 10 percent from adjacent DMAs
- The planned inclusion of the DMAs in any target audience emphasis plan
- Mail return rates from Census 2000
- Mail response in the American Community Survey

Early on in the planning process for this experiment, Census Bureau senior management expressed interest in selecting DMAs with a high concentration of HTC clusters. Consequently, our first criterion in narrowing potential sites was to isolate DMAs that skewed Economically Disadvantaged or Ethnic Enclave clusters[2]. The second consideration was size. Because of cost considerations, we mostly limited our selection to medium sized DMAs as increasing the paid media in larger markets would have been cost prohibitive. After carefully weighing the above criteria and consulting with the advertising vendor, we selected eight market pairs that would receive sufficiently similar treatments in sufficiently similar environments as to be considered evenly "matched". One DMA of each pair was randomly selected as the control site and the other was designated as the treatment site. The pairs are listed in Table 1 below. Appendix A contains a table illustrating selected match-pair criteria.

[1] These refer to clusters developed for the market segmentation for the 2010 Census Integrated Communication Campaign. In all, eight clusters were identified and each Census tract is assigned to one cluster. The clusters include: Advantaged Homeowners, All Around Average (owner skewed), All Around Average (renter skewed), Economically Disadvantaged (owner skewed), Economically Disadvantaged (renter skewed); Ethnic Enclave (owner skewed); Ethnic Enclave (renter skewed) and Single/Unattached/ Mobiles. For more information, see Bates and Mulry, (2011).

[2] Because we ran out of good match pairs that skewed economically disadvantaged and ethnic enclave, two of the market pairs selected skewed All Around Average cluster.

Table 1. Experimental DMA Pairings

DMA Segmentation Type and Size	Control		Treatment
Econ. Disadvantaged – medium	Columbia , SC	matched to →	Savannah, GA
Econ. Disadvantaged – medium	Toledo, OH	matched to →	Flint-Saginaw-Bay City, MI
Econ. Disadvantaged – medium	Jackson, MS	matched to →	Montgomery, AL
Econ. Disadvantaged – medium	Augusta, GA	matched to →	Baton Rouge, LA
Econ. Disadvantaged – medium	Tallahassee-Thomasville, FL	matched to →	Shreveport, LA
Ethnic Enclave – medium	Lubbock, TX	matched to →	Odessa-Midland, TX
All Around Average – medium	Joplin, MO	matched to →	Erie, PA
All Around Average – large	Little Rock-Pine Bluff, AR	matched to →	Jacksonville, FL

As part of the baseline advertising plan, four of the pairs were slated for inclusion in the Black Audience Plan (BAP). This was a targeted emphasis plan to reach DMAs with sizeable populations of Black adults 18 and older including African Americans, Afro-Caribbeans, Haitians and Black Africans. Media channels in the BAP included targeted syndicated television and national cable, local and syndicated radio, and community newspapers. So, DMAs slated for this plan received additional media reaching all segments of the Black Diaspora, in addition to the Diverse Mass (DM) audience advertising. Pairs that received the BAP included: Columbia-Savanna, Jackson-Montgomery; Augusta-Baton Rouge; and Tallahassee-Shreveport. None of the other pairs fell into any other targeted emphasis plan.

On average, the DMAs represented areas with a population aged 18 and older of around 500,000. The two exceptions were the Toledo-Flint pair which have populations 18 and older around 850,000 and the largest pair of Little Rock-Jacksonville are each over 1 million. Given the leadership input, matching criteria, and cost considerations that had to be factored when selecting pairs, we acknowledge they do not represent a random sample of all DMAs. Consequently, the inferences in this report cannot be generalized more widely to all DMAs.

2.2 Design of Experimental Treatment (Dosages)

The design for the experiment is shown below in Table 2.

Table 2. Design for Paid Media Heavy-up Experiment

Panel	Objective	Treatments
1	CONTROL	Control: Media level delivered through the integrated communication plan (ICP) national media base plans
2	TREATMENT	100 percent Heavy-up of Diverse Mass plan: 100 percent increase of media level in ICP national base plans delivered through local media buys.

The media strategy for the treatment sites was to use local media buys to increase the media level in the mass communication base plan by 100 percent. Industry media level tests typically use a 50 percent - 100 percent increase in media level. Lodish, et al. (1995) found an average increase of 85 percent for established brands over the 141 tests they analyzed. The media buys were designed to replicate the timing, media mix, and mix of specific messages from the national plan in the form of a "heavy-up" media schedule. The additional media ran during the awareness and motivation phases of the ICC with a goal of effectively doubling the media intensity in the treatment DMAs as compared to the control DMAs. The awareness phase occurred January 17 through February 28 while the motivation phase ran from March 1 through mid-April.

It is important to note that the extra media purchases were restricted only to advertising targets of the *DM campaign* – that is anyone who consumes English language media. The additional buys occurred in local television and radio broadcasts, print media, and online media (neither magazines nor out of home advertising such as billboards were part of the PAHUE). Local on-line media outlets contained content that was specifically designed to appeal to viewers/readers in the local market and was unique to that market.

2.3 Other Design Considerations

Early in the experimental design, the possibility of reducing the level of advertising in some DMAs was explored. This approach was rejected for several reasons for this test. First, the operational difficulties of "cutting-out" some or all of the network-delivered Census advertising in a DMA are quite high. While cut-outs can be accomplished, there are no good alternative uses for the advertising spots within the Census Bureau or the government without contaminating the test. The advertising time would have to be sold to other advertisers -- a difficult task due to the localized nature of the cut-outs. Second, the workload for nonresponse followup (NRFU) would be increased to an unknown extent. Third, advertising reductions might adversely affect morale and support for the census among partnership groups and local governments.

2.4 Rapid Response Spends

A critical part of the ICC involved an advertising budget reserve to make last minute ad buys in underachieving markets (referred to at the Rapid Response Program). These markets were identified by consulting daily mail participation rates. In designing our experiment, we assumed that no special interventions would occur in any of the treatment sites as a result of lower than anticipated mail response rates.

In mid-April we were notified that incremental Rapid Response radio and television media buys had been placed in five of the eight PAHUE pairs of sites. In two of the pairs (the Augusta-Baton

Rouge and Jacksonville-Little Rock pairs) the extra media buys were purchased in the *treatment* sites, therefore not compromising the integrity of the experiment. In the remaining three pairs (Flint-Toledo, Shreveport-Tallahassee and, Lubbock-Odessa), the extra buys occurred in the *control sites,* causing an imbalance in the heavy-up test. At the request of the Census Bureau, the advertising agency made additional last-minute radio and television buys in the *treatment sites* for these three pairs to try to recalibrate back to a doubling of media spend. These buys occurred very late in the ICC, just after the motivation phase concluded (during the week of April 15). In two of the sites, we were able to buy at least half the extra spends that occurred in the control (in Flint-Saginaw and Shreveport). However, due to market saturation, the agency was not able to purchase enough extra advertising in the remaining treatment site (Odessa) to completely rebalance that pair.

In addition to the incremental Rapid Response spends on spot TV and radio purchased by the agency, additional local buys were made under the supervision of the Regional Office Directors. These also included local spot TV, radio, and newspapers. Due to the timing of the buys, the localized nature of the Regional Director's buys, and inability to completely rebalance the pairs, the Rapid Response interventions undoubtedly introduced some element of contamination to the larger test. It also highlights a major "real life" challenge to implementing media experiments during a Census.

2.5 Targeted Audience Spends

In addition to the campaign for the DM audience, there were campaigns targeted at specific racial and ethnic audiences. A criterion used for identifying potential pairs of DMAs was that both members of a pair would either be included or excluded from any campaign aimed at a specific racial or ethnic audience. The BAP directed substantial, local spending to four of the eight PAHUE DMA pairs. BAP media buys were not intentionally varied between the treatment and control DMAs, and records show that media delivery was roughly equivalent *within* the pairs. There was also local spending targeted for the Hispanic audience. The level of spending aimed at the Hispanic audience in these DMAs was much lower than that aimed at the DM or Black audiences. This spending was again roughly equivalent between the members of the pairs. While the level of spending directed at particular racial and ethnic audiences was roughly equal between the members of a pair, this additional component tended to reduce the percentage difference in spending levels between treatment and control areas in the affected pairs.[3]

2.6 Actualized Heavy-Up Spends and Gross Rating Points

The total actualized buys for the heavy-up buys during the awareness and motivation phases was 1.24 million dollars. Approximately 70 percent was spent on spot television buys including local network affiliates and cable television. Another 13 percent was spent on local spot radio advertisements while 6 percent went toward local newspaper ads (which were run only during the motivation phase). The final 11 percent was spent on geographically targeted digital advertising such as internet banner ads.

[3]The heavy-up only manipulated spends in the Diverse Mass campaign and did not include additional buys in any of the targeted audience plans. Including the targeted plans in the experiment would have been difficult based on a lack of available targeted media inventory and the fact that targeted audience buys were made by several different minority advertising agencies subcontracted by DraftFCB.

We calculate total spending by adding the actualized spend from the heavy-up to amounts allocated to the sample DMAs from the national campaign spending, Rapid Response buys, and targeted audience buys to determine the total amount of spending in each market (shown in Table 3 for the DM campaign and in Table 4 for the total spending for all audiences)[4]. These figures also include NRFU spending which was about 7 percent of the total plan. NRFU buys cannot be separated from Awareness and Motivation phase buys in the local DMA post-buy spending reports. Of course, the level of spending in a market is positively correlated with the population of the market. Also, the basic cost of media varies from market-to-market depending on overall availability, demographics, geographic issues, and competition among media outlets. Looking at spending per adult is a quick way to adjust for differences in the level of advertising activity in a DMA due to market size. However, this measure is still subject to distortion from other factors listed above, such as overall availability and demographics.

Table 3. Diverse Mass Audience – Actualized Media Buys by DMA

DMA Name	Total Adults 18 and older in the TV DMA	Diverse Mass Media Spending by DMA (in $)	Percent Change from Control to Treatment DMA	Diverse Mass Media Spending per Adult 18 and older (in $)	Percent Change from Control to Treatment DMA
Columbia, SC	750,366	213,304	48	0.28	78
Savannah	622,608	314,890		0.51	
Toledo	833,051	224,233	79	0.27	71
Flint-Saginaw-Bay City	871,057	400,657		0.46	
Jackson, MS	644,554	196,370	53	0.30	120
Montgomery-Selma	446,644	299,564		0.67	
Augusta-Aiken	490,538	136,224	197	0.28	127
Baton Rouge	641,395	405,073		0.63	
Tallahassee-Thomasville	527,792	203,985	125	0.39	63
Shreveport	728,540	459,068		0.63	
Lubbock	301,938	93,438	93	0.31	106
Odessa-Midland	282,371	180,044		0.64	
Joplin-Pittsburg	294,496	83,346	131	0.28	125
Erie	302,032	192,353		0.64	
Little Rock-Pine Bluff	1,055,085	415,150	71	0.39	39
Jacksonville	1,299,788	709,521		0.55	
Total	**10,092,255**	**4,527,220**		**0.45**	

Source: DraftFCB database.

Table 3 illustrates the spend levels in the treatment and control pairs for the DM audience buys (the only media plan under control in the experiment). While the column labeled "Diverse Mass Spend by DMA" indicates the actual dollars spent, the column labeled "Diverse Mass Spend per A18 and older" is a more appropriate indicator of advertising intensity to reach the DM audience

[4] See Appendix B for detailed actualized local and national media spends by different media and audience plans.

– dollars spent relative to each person in the DMA aged 18 and older. Using this metric, the last column gives the percent increase in DM spending in the treatment sites relative to the controls. With the exception of the Little Rock/Jacksonville pair, the experiment was successful in obtaining at least a 50 percent increase for the DM media spend in the treatment markets. A 50 to 100 percent increase in media weight is an industry "rule of thumb" for the range of media increase where a change in the outcome variable (usually sales) should be observable in the test (see Lodish, et al. (1995).

Table 4 tells a less encouraging story. Here we show the *total* audience spends which add the targeted audience plan spends to the DM spending in Table 3. With this additional spending included, we see that the percent increase in the treatment sites fell short of a 50 percent increase in four of the pairs, and doubling occurred in only one pair (see last column in Table 4).

Table 4. Total Audience – Actualized Media Buys by DMA

DMA Name	Total Adults 18 and Older in the TV DMA	Total Audience Media Spending by DMA ($)	Percent Change from Control to Treatment DMA	Total Audience Media Spending per Adult 18 and Older ($)	Percent Change from Control to Treatment DMA
Columbia, SC	750,366	390,385	17	0.52	41
Savannah	622,608	457,970		0.74	
Toledo	833,051	328,400	49	0.39	42
Flint-Saginaw-Bay City	871,057	487,953		0.56	
Jackson, MS	644,554	398,083	9	0.62	57
Montgomery-Selma	446,644	433,343		0.97	
Augusta-Aiken	490,538	259,548	120	0.53	68
Baton Rouge	641,395	570,272		0.89	
Tallahassee-Thomasville	527,792	371,973	66	0.70	20
Shreveport	728,540	616,320		0.85	
Lubbock	301,938	185,042	41	0.61	51
Odessa-Midland	282,371	261,564		0.93	
Joplin-Pittsburg	294,496	89,476	122	0.30	117
Erie	302,032	198,917		0.66	
Little Rock-Pine Bluff	1,055,085	529,990	66	0.50	35
Jacksonville	1,299,788	882,134		0.68	
Total	**10,092,255**	**6,461,369**		**0.64**	

Source: DraftFCB database.

Another way to assess the level of per capita exposure to advertising in a market is through the rating points delivered during the campaign. The *rating* of an advertisement is the percentage of the target audience which had an opportunity to see or hear a given ad (i.e., tuned to the TV program, read the magazine, or listened to the radio at that moment). Each *rating point* corresponds to 1 percent of the target population having the opportunity to see an ad. The *Gross Rating Points* (GRPs) figure for a campaign is the sum of all of the ratings for all of the ads in the campaign. In practice, GRPs are measured only for television and radio. They can be estimated in advance for large circulation national magazines.

One way of interpreting the overall impact from a given level of GRPs on the target audience is in terms of the *reach and frequency* of the paid advertising campaign. The *reach* of a campaign is the percentage of the total target audience who were in the audience rating for at least one ad. The *frequency* of a campaign is the average number of times a target audience member who was reached by the campaign had the opportunity to be exposed to an advertisement. Frequency can be calculated by dividing the total GRPs by the percent of the population reached. Across all DMAs, the average DM GRP deliveries for the Awareness and Motivation phases (including Rapid Response) were 1,428, 2,004, and 354, respectively for a total of 3,786 GRPs. The estimated reach in the Awareness and Motivation phases was 97 percent. The Rapid Response phase was coincident with the Motivation phase and contributed to its 97 percent reach. This works out to the average person having about 39 potential exposures to the DM advertising during these two time periods.

Table 5. Diverse Mass Plan Gross Rating Points for Matched Pair DMAs

Control Site	Total GRPs	Treatment Site	Total GRPs	Percent Increase in Treatment DMA
Columbia	3,480	Savannah	6,128	76%
Toledo	3,504	Flint	6,663	90%
Jackson	3,597	Montgomery	6,091	69%
Augusta	3,663	Baton Rouge	6,158	68%
Tallahassee	3,601	Shreveport	5,850	62%
Lubbock	3,372	Odessa	5,937	76%
Joplin	3,442	Erie	6,131	78%
Little Rock	4,018	Jacksonville	7,203	79%

Source: DraftFCB database.

Table 5 contains the measured GRPs from the *Diverse Mass Plan* for adults 18 and older delivered by television and radio in the eight matched pairs. Despite the Rapid Response interventions and other differences between planned and actual spends, the experiment came close to achieving the goal of doubling the DM paid media GRP dosage in most of the treatment sites for the measured media of television and radio. For the DM plan, we see that GRP percentage increases in the treatment sites range from 62 to 90 percent.

However, when we consider the *total combined audience GRPs*[5] (which include DM rating points plus those associated with the targeted audience plans) differences between treatment and

[5] The GRPs delivered to a specific target audience from several different audience plans are the sum of the measured GRPs from each audience plan. The ratings for each target audience and audience plan can be obtained from national rating services, such as A.C. Nielsen. As an example, the national DM TV plan delivered a total of 2,322 TV GRPs to all adults 18 and older and 1,103 TV GRPs were delivered to Hispanic adults 18 and older. The national Hispanic audience plan delivered 2,473 TV GRPs to the Hispanic audience. Therefore, the TV GRPs delivered to the Hispanic audience through these two plans are 2,473 + 1,103 = 3,576 TV GRPs at the national level. The same procedure is applied across all measured media to get a value to total GRPs for an audience.

To obtain total GRPs for a geographic area, one combines the GRP values for the various audiences into a weighted sum where the weights are the proportion that a given audience is of the total population in the geographic area. For example, in the Columbia, SC DMA, the Black, Hispanic, and Non-Hispanic/Non-Black audiences are 36 percent, 3 percent, and 61percent of the TV audience, respectively. The GRPs for Awareness and Motivation phases were

control GRPs are greatly reduced in the pairs affected by targeted audience plans (see Table 6 and Figure 1). The reduced difference in GRPs between treatment and control sites is even more evident when the Black and Hispanic audiences are broken out separately (Tables 7 and 8). For example, the GRPs delivered to the Black audience include the DM media and also the BAP in the four pairs that received it. The "watering-down" of the heavy-up dosage is very evident in these pairs (i.e., Columbia-Savannah; Jackson-Montgomery; Augusta-Baton Rouge; and Tallahassee-Shreveport). The GRPs delivered to the Hispanic audience in the PAHUE markets come mainly from the national Hispanic plan which was delivered in *all* DMAs and did not vary across the PAHUE markets. This, combined with the fact that Hispanics consume a lower rate of DM media, meant the lion's share of Hispanic GRPs came from the national Hispanic plan as opposed to the DM plan. Thus doubling local DM buys in the heavy up had small effects on Hispanic GRPs[6]. Thus, like the Black audience, the increase in the treatment dosage for the Hispanic audience fell short of even a 50 percent increase in all but one pair (see Table 8, last column).

Table 6. Total Combined Audience Gross Rating Points for Matched Pair DMAs

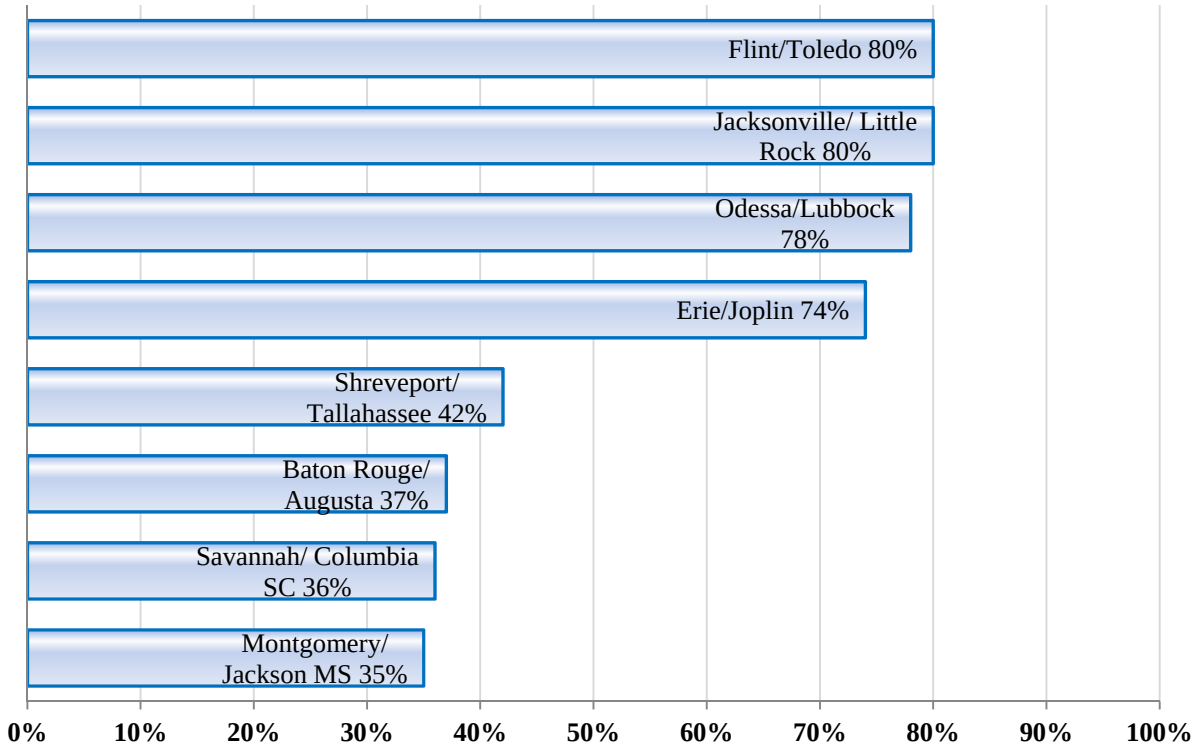
Control Site	Total GRPs	Treatment Site	Total GRPs	Percent Increase in Treatment DMA
Columbia	5,664	Savannah	7,684	36%
Toledo	3,994	Flint	7,171	80%
Jackson	6,590	Montgomery	8,895	35%
Augusta	5,856	Baton Rouge	8,013	37%
Tallahassee	5,448	Shreveport	7,718	42%
Lubbock	4,092	Odessa	7,285	78%
Joplin	3,613	Erie	6,301	74%
Little Rock	4,729	Jacksonville	8,490	80%

Source: DraftFCB database.

9,215, 4,763, and 3,613 respectively for each audience. The combined total GRPs for the Awareness and Motivation phases in Columbia were: (9,215 * 0.36) + (4,763 * 0.03) + (3,613 * 0.61) = 5,664 GRPs, as shown in table 6.

[6] The Odessa-Midland/Lubbock pair is the exception because these areas contain sufficient Hispanic populations for Nielson to estimate the Hispanic ratings. Lubbock has relatively low ratings for Spanish language programming, diluting the impact of the national Hispanic plan and thereby increasing the heavy-up lift seen in Odessa-Midland.

Figure 1. DMA Pairs Ordered by Heavy-Up Increase in GRPs



Source: DraftFCB

Table 7. Black Audience: Total Combined Audience Plans Gross Ratings Points for Matched Pair DMAs

Control Site	Total GRPs	Treatment Site	Total GRPs	Percent Increase in Treatment DMA
Columbia	9,215	Savannah	10,100	10%
Toledo	7,188	Flint	10,184	42%
Jackson	9,226	Montgomery	11,763	28%
Augusta	8,768	Baton Rouge	10,480	20%
Tallahassee	8,555	Shreveport	10,946	28%
Lubbock	5,797	Odessa	8,676	50%
Joplin	6,099	Erie	8,132	33%
Little Rock	6,390	Jacksonville	11,639	82%

Source: DraftFCB database.

Table 8. Hispanic Audience: Total Combined Audience Plans Gross Ratings Points for Matched Pair DMAs

Control Site	Total GRPs	Treatment Site	Total GRPs	Percent Increase in Treatment DMA
Columbia	4,763	Savannah	6,046	27%
Toledo	4,773	Flint	6,311	32%
Jackson	4,818	Montgomery	6,032	25%
Augusta	4,849	Baton Rouge	6,065	25%
Tallahassee	4,822	Shreveport	5,921	23%
Lubbock	3,498	Odessa	5,903	69%
Joplin	4,744	Erie	6,049	27%
Little Rock	5,015	Jacksonville	6,539	30%

Source: DraftFCB database.

Using the total GRPs combined across audience plans as in Table 6, we can observe the frequency with which the DMA audience had the opportunity to see Census Bureau paid advertising. For example, in Table 6 the total combined audience GRPs in the control site Toledo is 3994. So, the estimated average frequency of exposure to a Census television or radio ad between January 11 and April 18 for an adult in Toledo aged 18 and older was 41.18 times (3,994 GRPs / 97 percent reach = 41.18). This compares to 73.93 times in the paired treatment site of Flint (7,171 GRPs / 97 percent reach = 73.93). When we look at the black audience reach and frequency, some of the numbers are staggering. For the same pair, we see the estimated frequency of ad exposure for the Black population in the *control* site of Toledo was at 74.10 (7,188 GRPs / 97 percent reach = 74.10) while the Flint treatment site was 104.99 (10,184 GRPs / 97 percent reach = 104.99).

These numbers give a sense of the breadth and depth of the campaign for some audiences. Even in the control sites, the media spend levels were among the top five heaviest advertising campaigns during the time period. Between January and June 2010, the Census Bureau ranked fourth out of the top six advertisers behind McDonalds, Wal-Mart and Geico but ahead of Budweiser and Nike. Knowing this provides context as to whether the control sites may have been saturated – given such a large control dosage, more advertising may have little effect. That said, we point out the one pair, Joplin-Erie, that came closest to achieving the experimental dosage difference originally planned for the test. At least from a total audience standpoint, this pair had a close-to-average control GRP dosage (3613 DM GRPs versus the national level of 3786 DM GRPs) and was successful in delivering a 74 percent heavy-up of additional GRPs to the treatment site (as shown in Table 6).

2.7 Data Sources

We make use of two data sources to examine outcomes in the treatment and control areas. The first is survey data on census knowledge, attitudes, and advertising recall that were collected from a sample of respondents in the PAHUE DMAs. We use these data to examine whether the increased advertising in treatment sites is reflected in greater ad recall, in shifts in knowledge and attitudes towards Census, or in mail return behavior.

The second source is operational data from the Census for all housing units in the PAHUE DMAs that were eligible to mail back a census form. This provides information on the primary outcome of interest—mail return rates—along with factors that might influence rates of return such as household composition, housing characteristics, and whether a replacement form was sent.

2.7.1 Survey Data

As part of the effort to evaluate the 2010 Census ICC, the Census Bureau contracted with the National Opinion Research Center (NORC) to design and conduct an independent and comprehensive evaluation of the 2010 Census ICC. NORC conducted a set of three household surveys based on a nationally representative sample and oversampling of minority populations and other targeted segments. This data collection is known as the 2010 Census Integrated Communication Program Evaluation (CICPE). The survey items included measures such as self-reported Census ICC exposure, Census awareness, knowledge, and attitudes and self-reported intent to participate. The first wave was conducted prior to the paid advertising to assess baseline levels of Census awareness, knowledge and intent to participate (mid-September to mid-January, 2010). The second wave took place January 19 through March 18, 2010 during the peak of the paid media campaign. The third wave was conducted mid-April through mid-July 2010 when door-to-door enumerators conducted nonresponse followup interviews for those households failing to complete and return a census form. This wave contained an additional measure of campaign fatigue that might be associated with the extra dose of advertising.

In order to assess 2010 Census knowledge, attitudes and advertising recall in the PAHUE sites, we had NORC add sample to the CICPE drawn from the PAHUE DMAs. NORC selected census tracts within each of the PAHUE DMAs. To control travel costs, NORC mostly sampled from counties located within a 40 mile radius of the DMA central city. Additionally, some counties were taken out of sample to avoid interviewing in areas at risk of "spill in" advertising from adjacent DMAs. These decisions effectively removed 90 counties of the total 267 involved in the experiment. Within the remaining 177 counties, households were selected with equal probability within each DMA.

NORC used an address based sampling strategy already being used by the larger CICPE data collection. This involved matching phone numbers with sampled addresses, contacting households with associated phone numbers in a Computer Assisted Telephone Interview CATI environment, sub-sampling non-CATI completed households (as well as those addresses that could not be matched to telephone numbers) and fielding the subsample using face-to-face interviewing. The PAHUE interviews were collected during the CICPE wave 1 and wave 3 field periods. This provided both pre and post heavy-up intervention interviews. Table 9 illustrates the PAHUE response rates and completed interviews by mode. Response rates were calculated using the AAPOR RR2 formula (see AAPOR, 2011).

Table 9. PAHUE Survey Interviews

	Wave 1	Wave 3
CATI interviews	808	844
Face to face interviews	1,116	1,202
Total	**1,924**	**2,046**
Weighted Response Rate	68.3%	70.8%

Source: Waves 1 and 3 of CICPE survey.

Readers can find more information about CICPE data collection (including an overview of survey data items being collected) by consulting the 2010 Census ICP Evaluation (see Datta et al., 2012).

2.7.2 Operational Data

Our use of Census operational data was made possible by using an extract of response data prepared for another 2010 evaluation – the 2010 Mail Response/Return Rates Assessment (see Letourneau, 2011). This extract contained the following information:

- variables needed to determine mail return status,
- the dates on which returned forms were received,
- household information collected as part of the 2010 Census (e.g., age, sex, race/ethnicity of householder, number of children, whether unit was owned or rented),
- whether the household received a replacement and/or a bilingual form, and
- type of housing unit.

These variables, along with the level of paid advertising "dosages," were used to analyze the determinants of the share of households that completed and mailed back a census form by April 19, 2010 (the cutoff date for sending cases on to personal visit NRFU). For our report, we followed the same guidelines for calculating mail return rates used in the 2010 Mail Response/Return Rates Assessment. For housing units with no mail return, the extract also included information on whether the outcome of NRFU operations was a direct response from a household member or if only a proxy response was obtained.

Note that we also have information on household mail return status for the CICPE PAHUE samples. This information was merged in by the Census Bureau by matching the CICPE sample addresses to the Master Address File (MAF) and then extracting Census operational variables needed to determine if a household mailed back a form prior to nonresponse followup. This information allows us to analyze the survey data on knowledge, attitudes, and ad recall in conjunction with actual mailback behavior. Details of this extract can be found in Datta et al., 2012. While this allows us to estimate treatment/control differences in return behavior from the CICPE sample, we have information for all eligible residents rather than a sample from the operational data base, so that is the preferred source for comparisons of rates of response.

2.8 Analytic Approach

To assess the impact of the heavy-up increase in media spending, we carry out several types of analyses. We discuss the details of these analyses in the next sections, but here we outline our approach and the reasoning behind it.

The primary goals of the 2010 ICC were to increase public awareness and positive attitudes towards the census and, in doing so, to increase the number of households returning a completed census form before costly followup operations began. With this in mind, we first evaluate the effects of the heavy-up treatment on intermediate outcomes—awareness and attitudes—in treatment versus control areas using the CICPE data. The specific outcomes that we consider are self-reported exposure, recall of specific components of the campaign, and attitudes towards the census. We then turn to the operational data to examine whether treatment sites differed from controls in their rates of pre-NRFU mail return, in the timing of those returns, or in the share of NRFU returns that took the form of proxy response.

The treatment versus control comparisons take two forms. The most common is simply to estimate whether treatment sites on average have more positive outcomes. But with measures that were included in both waves of the CICPE survey, we can additionally analyze whether changes over time are more positive for treatment sites than control sites—an approach sometimes referred to as differences in differences. Because additional advertising might have larger effects on some types of households than others, where practical we carry out the same analyses separately for different race/ethnicity groups and for different campaign audience segments. Assignment to a particular audience segment is based on characteristics of the tract of residence rather than on the characteristics of individual households.

For much of the analysis, the statistical methods used are simply tests for differences in means. While DMAs within a pair were matched on a number of characteristics to minimize differences in response rates in the absence of any experimental treatment, none of the pairs is made up of two identical DMAs. Given that we only have eight pairs of DMAs, effects of the experiment have to be reasonably large relative to inherent variation between paired DMAs or we will be unable to distinguish experimental effects from random differences. One way to improve our ability to make this distinction is to adjust for any obvious differences using the information we have on ways in which DMAs differ within pairs—for example, differences in the demographic characteristics of their populations. So in addition to doing simple treatment/control comparisons of means, we also use logistic regression models to account for measurable differences across areas that are associated with different levels of mail response. Using multivariate models also allows us to incorporate information on the GRPs achieved in each DMA to more carefully calibrate the size of the experimental treatment.

The final part of the analysis is a cost-benefit component to judge the tradeoffs between the increased investment in media and the impact on mail return with its decrease in enumeration costs. Analysts expect to see variations in the impacts by DMA pair and plan to examine difference between pairs for significance. This component will have limitations based on projected media costs and assumptions about per household cost of conducting personal-visit nonresponse followups.

2.9 Variance Estimation

In each of the statistical analyses, we use design-based estimators that account for stratification and clustering, and in the case of the CICPE data, varying probabilities of selection. We treat the sample design as one in which each member of a pair of DMAs represents a randomly selected unit, and both members of the pair are drawn from the same stratum. In fact, DMA pairs were selected purposively rather than using probability sampling methods, and DMAs were paired before selection rather than independently selected. However, we think that the assumptions we use provide a reasonable approximation.

Within DMAs, the CICPE sampling is based on standard methods, and in using these data we apply weights that NORC developed to account for nonresponse and for sub-sampling of non-CATI-completed households. Because the operational data include all eligible households, the probability of selection is one for each housing unit, and weights are unnecessary.

The first two analytic methods in the plan above (analysis at the DMA level and at the tract within DMA level) will use linear models to determine the impact on the MRR from changes in paid media. The standard error for the effect of an increase in paid media will be estimated through the mean squares in the analysis of variance table for the model. It is expected that variation in the DMA MRRs will impact the size of the mean square used in the statistical tests. The contribution from DMA MRR variation will be reduced through the pairing of treatment and control DMAs.

The third method of analysis will look at the effect on the probability of response at the individual household level due to a change in paid media. The statistical analysis will use logistic regressions to estimate the dosage impact on probability of response. The standard error for this effect will be estimated through the regression model.

The final method will look at household survey data responses before and after the heavy-up intervention. Pairwise comparisons will be used to test differences between the combined treatment and control sites. Analysis will estimate standard errors via a Taylor Series approach to adjust for the clustered sample design.

3. Limitations

3.1 Challenges and Risks to Implementing a Heavy-Up Test

There was some risk that the additional media buys would not be executed exactly according to the heavy-up media plan. This sometimes happens when the desired media buys at a given time slot, in a particular medium, or in a geographic area are not available. This could result in a less than anticipated “dosage” increase in the treatment sites. However, for the DM media buys under the control of our test, this risk was not observed. There was also a risk that the individual sites within the matched pairs would differ substantially in the level of partnership or public relations intensity due to local governmental interest in the Census, availability/interest of local partners, and interest from local news outlets. The 2010 Census ICC included a number of inter-connected elements including paid advertising, earned media, promotional events, and a nationwide community-based partnership program. The PAHUE manipulated only one piece of one component.

There was also risk that special interventions might occur in one or more of the treatment sites to correct for lower than anticipated mail response. For example, several of the control sites exhibited lower than anticipated response in the early days of the Response Rate Feedback Program. As a result, decision makers intervened to correct the site's course. This contaminated the test to some extent (see Section 2.4).

Finally, there was a risk that one or more of the treatment sites would be at or near saturation levels of paid advertising. Heavy-up tests have difficulty detecting an effect from increased media investment if the base plan is near saturation. As noted in section 2.6, we believe the test suffered this risk to some degree.

3.2 Other Limitations of the Study

The PAHUE was conducted using mid-sized DMAs, which did not receive increased local DM media over and above the national plan. Some of the treatment-control pairs received an increase in targeted media for specific ethnic or cultural audiences; however, these increases were matched within the pairs. Mid-sized DMAs represent about 35 percent of the U.S. population and average about 250,000 households per DMA. The market pairs selected for our test average about 275,000 households per DMA. Consequently, they are not representative of the top 25-50 media markets. However, the results of the test could be used to estimate the impact from the mass communication base plan to the target populations in mid and smaller sized DMAs. With some care, the tract-cluster results may also generalize to similar tract-clusters for outlying counties of relatively large DMAs. The test can clearly guide judgment when estimating the effectiveness of 2010 paid media across sections of the U.S., but should not be used as a national assessment of the paid media plan.

Because the experiment required us to match DMA pairs according to a variety of selection criteria, the use of probability sampling was not appropriate to select the areas included in the experiment.

There is one main question under test. That is, would an increase in the DM plan be cost effective in terms of increasing MRR for each target audience (Asian, Black, Hispanic, non-Hispanic White). If no statistically significant differences are found, there are several potential interpretations. Lack of a significant treatment-control difference does not necessarily indicate that the advertising was ineffective. If the messaging appears to have the ability to change awareness and attitudes toward the Census in the pre/post surveys, we might suspect that the paid media was near saturation levels for the non-Hispanic White audience, especially in cooperative tract-cluster types, like Advantaged Homeowners. With advertising near saturation, it would be appropriate to test reductions of paid media levels in a future test. For the Asian, Black, and Hispanic target audiences, we might suspect that mainstream media alone might not be sufficient to increase MRR. If there are large DMA-to-DMA differences in MRR within the treatment and control panels, we may need to investigate whether covariates that were not well measured (perhaps differing magnitude in foreign-born population or level of support from partners) may have added noise to the DMA outcomes, reducing the ability to detect significant differences in MRR associated with the treatment.

Some of the test's results could provide input to the media planning for the 2020 Census. Understanding the degree to which differences in the DM plan affected MRR in both ethnic and non-ethnic markets would be useful when planning the mix of mass and ethnically targeted communication vehicles in 2020. The heavy-up test was run with spending levels that were near the original budget levels for the 2010 Census, i.e., prior to the stimulus act increase[7]. Understanding how much MRR impact was produced at that spending level could help inform the initial budget level planning for the 2020 Census.

A final limitation involves the household level pre and post-intervention interviews. As noted, the PAHUE sample achieved a response rate of 68 percent and 71 percent for waves 1 and 3, respectively. While these levels of nonresponse are not too concerning, it is reasonable to believe that non-respondents may be different in their attitudes, knowledge and behavior from those that did participate. If this is the case, then our outcomes of analytical interest may be subject to nonresponse bias. Based on analysis conducted by NORC, we believe this bias is close to zero for the PAHUE sample (see Datta, et al., 2011 for further detail). However, this possibility is one reason to favor mail return results based on the operational data rather than the CICPE data, as it is not subject to nonresponse bias. Finally, due to cost and resource constraints, the ICC evaluation team agreed to limit the geographic scope of the PAHUE interviews. With few exceptions, sample cases were limited to households located in counties within a forty-mile radius of the DMA primary central city. A few other counties were eliminated due to potential media spill-in from neighboring DMAs.

4. Results

4.1 Self-Reported Ad Exposure

One basic question about the effects of the PAHUE is how much the additional advertising in heavy-up areas increased exposure to advertising among the treated population. We know from the ratings data (GRPs) that there were substantial increases in actual exposure to the ads. The CICPE survey gives us one way to measure this by examining treatment-control differences in self reported exposure. In both waves 1 and 3, respondents were asked a series of questions about whether they recalled hearing or seeing information about the census. The survey asked separately about exposure via different media outlets, including paid advertising, Census's community partners, and various sources of news coverage. Because the heavy-up increase in media buys occurred after wave 1 data were collected, the experiment could not have had a differential effect on media exposure for the treatment areas in wave 1. Wave 3 data were collected after the heavy-up treatment had taken place, so we would expect to see its effects in evidence in estimates for that wave.

We note, however, that self-reports of media exposure are prone to measurement error and far from a perfect indicator. For example, in their investigation of error in self-reported exposure to 2010 Census paid advertising, NORC found confirmed awareness to be low relative to self-reported recall. That is, the majority of respondents who claimed to see a particular ad failed to recall meaningful details to confirm awareness. Further, the degree of confirmed awareness was

[7] As part of the American Recovery and Reinvestment Act (ARRA), the 2010 Census campaign received additional funding in August 2009 to spend on advertising and promotion.

found to vary by age, education, and media outlet (Datta et al, 2011). This is a limitation to be kept in mind throughout the section.

Tables 10 and 11 present measures of self-reported exposure to paid advertising. The difference between the two tables is simply the way in which we quantify exposure, though in the end both measures produce similar evidence. These measures are based on the following questions from the CICPE survey, where question 18 was asked only if the respondent said yes to one of the parts of question 17:

Q17. Have you heard or seen advertisements about the census...
- a. on television?
- b. on the radio?
- c. in magazines?
- d. in newspapers?
- e. on the Internet?
- f. in other places such as coffee cups, billboards, or park benches?

Q18. Thinking about all of the advertisements you heard or saw in the **past 30 days** about the census, how many different times in the **past 30 days** would you say you saw or heard something about the census? [Note that the wave 3 questionnaire asked about exposure in the **past 90 days**.]

Response categories:
- None
- Once or twice
- 3-5 times (W1) / 3-15 times (W3)
- 6-10 times (W1) / 16-30 times (W3)
- 11 times or more (W1) / 31 times of more (W3)

Table 10 presents estimates of whether respondents reported having seen any ads about the census, while in Table 11, we compare differences in the number of times respondents reported seeing or hearing a census ad. Throughout this section, three tests of differences are typically presented. The first test is the difference between the wave 1 treatment and wave 1 control. This provides a benchmark of how close the two measures were prior to advertising. The second is the difference between wave 1 and wave 3 for the control cases. This is a good indicator to measure change over time as a result of *normal* campaign levels. It also serves as an indicator of potential saturation as a result of normal campaign dosages – that is, if wave 3 levels for the control are very high, we might conclude the selected markets were at or near the saturation point. The final test is whether the *difference* between the treatment and control wave 1-wave 3 differences is significant – this is the test of most interest because it tells us whether differences between the treatment and control were large enough to be statistically meaningful. Each of the tests presented in the table is based on a t-test of the hypothesis that differences equal zero.[8]

[8] The standard error used to calculate the t-statistic is estimated using Taylor series (or linearization) methods that adjust for the complex design of the CICPE survey

Table 10. Any Exposure to Paid Advertising About Census, Treatment Versus Control

	Percent Heard or Saw at Least One Census Ad	
	Wave 1	Wave 3
Treatment	30.4	76.0
	(2.6)	(2.0)
Control	37.7	76.5
	(3.5)	(4.4)

Test	Estimate
Treatment minus control in wave 1	-7.3
	(5.5)
Change for control (wave 3 minus wave 1)	38.7***
	(3.2)
Change in treatment minus change in control	6.9
	(5.4)

p-values: *<.10, **<.05, ***<.01
Base: Weighted estimates using all respondents. Respondents who reported they had never heard of the census were coded as having heard or seen no ads.
Source: Waves 1 and 3 of CICPE survey.

In both Tables 10 and 11, there are large and significant increases in reported media exposure between waves 1 and 3, but the increase in treatment areas is not significantly different from the increase in control areas.

Table 11. Frequency of Exposure to Paid Media, Treatment Versus Control

	Number of Times Heard or Saw Paid Ad in Last…	
	Wave 1 (Last 30 days)	Wave 3 (Last 90 days)
Treatment	0.81	10.37
	(.08)	(0.97)
Control	1.18	9.50
	(0.11)	(0.80)

Test	Estimate
Treatment minus control in wave 1	-0.37**
	(0.13)
Change for control (wave 3 minus wave 1)	8.32***
	(0.73)
Change in treatment minus change in control	1.24
	(1.32)

p-values: *<.10, **<.05, ***<.01
Base: Weighted estimates using all respondents. Respondents who reported they had never heard of the census were coded as having heard or seen no ads. Responses to question 18 were given in intervals (0, 1-2,etc). Each respondent was assigned the mid-point of their interval response (e.g., 1.5 for the category 1-2 times). For wave 1 we assigned a value of 15 for the category 11 or more times and for wave 3 we assigned a value of 40 for the category 31 or more times. This is consistent with methods used in Datta, Yan, et al (2011)
Source: Waves 1 and 3 of CICPE survey.

Table 12 presents estimates of exposure for the specific types of media outlets listed under question 17. Only TV and radio were included in the heavy-up treatment, so we would expect to see evidence of increased exposure only for those media. Like the results above, these estimates consistently show significant increases in exposure between waves 1 and 3. The increases in

reported exposure in control areas are significantly positive for all types of media, with television showing the largest change. Here we find some evidence that the increase in reported exposure was larger for the treatment than the control group. The treatment group reported a larger increase in exposure to radio ads than the control group, though this difference is significant only at the 10 percent level. However we do not find a significantly larger increase in TV exposure. Finally, we also see a marginally significant difference in reported magazine exposure between treatment and control over waves with the control sites reporting a larger difference. Since magazine ads were not manipulated as part of the heavy-up buy, we do not consider this finding relevant to our experiment.

Table 12. Exposure by Type of Media, Treatment Versus Control

	Percent With Exposure Through Media of This Type											
	TV		Radio		Newspapers		Magazines		Internet		Other	
	W1	W3	W1	W3	W1	W3	W1	W3	W1	W3	W1	W3
Treatment	20.3 (2.3)	68.4 (1.9)	7.6 (1.0)	40.3 (1.8)	15.6 (1.5)	28.5 (1.7)	6.9 (1.3)	12.0 (1.4)	11.0 (1.1)	17.6 (1.5)	4.8 (0.7)	27.9 (3.2)
Control	26.5 (3.3)	69.3 (3.6)	12.8 (1.4)	37.7 (3.0)	18.4 (1.6)	34.1 (2.3)	5.4 (0.7)	15.2 (1.7)	14.7 (2.0)	24.3 (1.6)	6.0 (0.7)	36.1 (3.5)

Test	**Estimate**	**Estimate**	**Estimate**	**Estimate**	**Estimate**	**Estimate**
Treatment minus control in wave 1	-6.2 (5.4)	-5.2** (2.1)	-2.8 (2.7)	1.5 (1.6)	-3.8 (2.8)	-1.2 (0.8)
Change for control (wave 3 - wave 1)	42.8*** (3.9)	24.9*** (2.5)	15.7*** (1.8)	9.8*** (2.0)	9.5*** (1.3)	30.2*** (3.0)
Change in treatment minus change in control	5.3 (5.4)	7.7* (3.8)	-2.8 (3.6)	-4.7* (2.1)	-2.9 (2.4)	-7.0 (5.4)

p-values: *<.10, **<.05, ***<.01
Base: Weighted estimates. Excludes refusals. Those responding "Don't know" or who reported they had never heard of the census were coded as having zero exposure through each type of media.
Source: Waves 1 and 3 of CICPE survey.

Assuming that respondents had very precise recall of where they heard about the census, we would expect to see heavy-up effects only in measures of paid advertising exposure. But with imperfect recall, the effects might be better captured in a broader measure of exposure. With this in mind, in Table 13 we present estimates of combined rates of exposure to paid and earned

media, where the latter includes stories or features about the census in news outlets. Earned media exposure was elicited using questions similar to those for paid media:

> Q24. Not including advertisements, have you heard or seen any stories or features about the census…
>
> a. in a newspaper or magazine article?
> b. on television or radio?
> c. on the internet?
>
> *If yes to 24c, please answer the following…*
>
> On the internet, did you hear or see anything about the census on…
>
> d. internet blogs?
> e. social networking sites (e.g., Facebook, Myspace)?
> f. regular web sites?
>
> 25. Thinking about all of the places you heard or saw stories or features about the census in the **past 30 days**, how many different times in the past 30 days would you say you heard or saw something about the census? [Note that the wave 3 questionnaire asked about exposure in the **past 90 days**.]

We combine the measures of paid and earned media exposure by creating a variable indicating whether or not a household had above-median reported exposure to both types of media. Table 13 also indicates that the increase in exposure between waves 1 and 3 was significant, but again, we fail to reject the hypothesis that that increase was the same for treatment and control groups.

Table 13. Combined Measure of Exposure to Paid and Earned Media, Treatment Versus Control

	Percent With Above Median Exposure to Both Paid and Earned Media	
	Wave 1 (Last 30 days)	Wave 3 (Last 90 days)
Treatment	17.7	33.6
	(1.3)	(2.5)
Control	20.8	41.7
	(2.5)	(3.8)

Test	**Estimate**
Treatment minus control in wave 1	-3.1
	(3.4)
Change for control (wave 3 minus wave 1)	20.9***
	(4.4)
Change in treatment minus change in control	-5.0
	(5.4)

p-values: *<.10, **<.05, ***<.01

Base: Weighted estimates using all respondents. Respondents who reported they had never heard of the census were coded as having no exposure. In wave 1, the median level of exposure was zero for both paid and earned media, so high exposure is defined as some exposure to both types of media in the last 30 days. In wave 3, the median was 0 for earned media and having seen 1-2 ads for paid media, so high exposure is defined as some exposure to earned media and having seen at least 3 ads in the last 90 days. Source: Waves 1 and 3 of CICPE survey.

Table 14 reports an even broader measure of exposure that includes all sources. The measure here is based on the question:

Q8. Have you heard or seen anything recently about the 2010 Census?

For this measure, there is a large difference between estimates for treatment and control in wave 1, with the control group having a much larger estimated level of exposure. In wave 3, roughly 80 percent of both groups report having heard or seen something recently—a substantial increase from reports in wave 1. The bottom part of the table shows that the increase in exposure is significantly different from zero, but the wave 1 difference between treatment and control samples is not. While the estimate of exposure for the treatment group increased more between waves 1 and 3 than the estimate for the control group, this treatment/control difference in the change is not significantly different from zero either.

Table 14. Any Recent Exposure to Messages About Census, Treatment Versus Control

	Percent Heard or Saw Something About Census Recently	
	Wave 1	Wave 3
Treatment	43.4	80.5
	(3.9)	(2.6)
Control	56.7	79.9
	(6.7)	(4.6)

Test	Estimate
Treatment minus control in wave 1	-13.4
	(8.8)
Change for control (wave 3 minus wave 1)	23.2***
	(5.7)
Change in treatment minus change in control	14.0
	(7.8)

p-values: *<.10, **<.05, ***<.01
Base: Weighted estimates using all respondents. Respondents who reported they had never heard of the census were coded as not having heard or seen anything recently.
Source: Waves 1 and 3 of CICPE survey.

Thus this evidence from the CICPE survey consistently indicates that reported exposure to messages about the census increased between waves 1 and 3. While the treatment/control differences in changes are often in the right direction, we only find one measure with a significantly larger increase for the treatment group than the control group: self-reported exposure to radio advertising.

One possible reason we find so little evidence of a response to the heavy-up treatment here is that census advertising in non-treated areas was so high that the heavy-up increase occurred in a range of diminishing returns. In other words, the control areas were already saturated with ads to the point that the extra dosage delivered by the heavy-up was impossible to detect using measures of self-reported ad exposure. Wave 3 of the CICPE survey includes a question which we use in Table 15 to examine this possibility:

> 19H1. Thinking of all the ads you have seen or heard about the Census, would you say there have been:
>
> a. Far too few ads
> b. Too few ads
> c. About the right amount of ads
> d. Too many ads
> e. Far too many ads

There are several ways to interpret results from this item. Reports of too few ads might indicate the campaign did not resonate or make much of an impression whereas too many ads might suggest saturation and fatigue. In the table, we combine the "Far too few" and "Too few" categories, and similarly the "Far too many" and "Too many" categories. The first column of estimates shows that respondents in treated areas are significantly less likely to say that they thought there were too few ads for census – a measure perhaps consonant with "feeling" an increase in the treatment areas. Additionally, the treatment areas were significantly more likely

to report that the number of ads were "about right" compared to the control areas. Lastly, we do not find a significant treatment/control difference in the share saying there were too many ads -- this evidence at least suggests the increased dosage was not perceived as an irritant.

Table 15. Paid Media Saturation, Treatment Versus Control

	Percent Saying There Are		
	Too few ads	About the right number of ads	Too many ads
Treatment	13.4	73.7	12.9
	(0.9)	(1.7)	(2.0)
Control	22.0	64.7	13.3
	(1.6)	(2.4)	(1.5)
Test of treatment minus control	-8.5***	7.7*	-0.5
	(1.9)	(3.4)	(3.1)

p-values: *<.10, **<.05, ***<.01
Base: Weighted estimates. Excludes those are have not heard/seen anything about census recently.
Source: Wave 3 of CICPE survey.

A final comment on why we failed to detect higher self-reported exposure in the treatment areas may be related to the small differences in GRP delivery noted earlier in Table 6. Recall that once the *total audience* GRPs were taken into account, the treatment dosages fell fall short of a doubling heavy-up.

4.2 Modeling Self-Reported Ad Exposure

The lack of evidence for increased exposure in treatment sites is puzzling, so here we take the opportunity to examine more closely the relationship between self-reported campaign exposure and the heavy-up experiment. In the PAHUE, we know which areas received a higher delivery of GRPs, and all other things being equal, would expect this to correlate with higher self reported campaign exposure. Despite the shortcomings of the heavy-up dosages, the PAHUE still affords a unique opportunity to explore the metric used widely throughout this evaluation and the larger CICPE evaluation conducted by NORC, that is, self-reported exposure to the 2010 Census campaign. The NORC evaluation relies heavily on this measure as a key independent variable to predict propensity to mail back a census form. The main hypothesis is that those with higher reports of self-reported exposure would be more likely to return the form.

Thus far, we have found little evidence to support this hypothesis. Does this fact call into question the validity and usefulness of self-reported exposure as an analytic variable for evaluation purposes? We cannot answer definitely because of complicating factors. For one, the PAHUE was not able to double up the advertising dosages as planned, and for another, the advertising dosage in the control sites may have been at or near saturation levels making it very hard to detect a "signal" in the treatment sites. Nonetheless, we take a last look at the accuracy of these self-reports by modeling self-reported exposure levels controlling for media consumption, demographic and lifestyle variables, and whether respondents resided in the treatment or control sites.

Table 16 presents results from an ordered logit model using wave 3 data to predict the number of times respondents reported seeing or hearing a Census paid ad in the last 90 days. Dependent variable categories consisted of: zero times, 1-2 times, 3-15 times, 16-30 times and 31+ times.

Table 16: Ordered Logit Model of Self-Reported Exposure to the 2010 Census Paid Media Campaign

Indep. Variable	Odds ratio	95% Conf Interval	p-value
Treatment	1.109	0.670 – 1.834	.649
Age: 18-29	---	----	----
30-44	1.056	0.749 – 1.488	0.725
45-64	0.620	0.344 – 1.117	0.098
65+	0.481	0.283 – 0.817	0.013
Female	0.985	0.794 – 1.223	0.878
Education:			
< High School	---	----	----
High School	1.606	0.952 – 2.709	0.070
Some college	1.911	1.196 – 3.054	0.013
College graduate	2.155	1.208 – 3.843	0.016
> College	2.039	1.112 – 3.738	0.027
Hours work per week:			
Zero	---	----	----
1-19	1.456	0.546 – 3.883	0.403
20-39	1.045	0.567 – 1.927	0.873
40+	1.184	0.994 – 1.410	0.057
Daily hours watch TV			
Zero	---	----	----
1-2	2.582	1.567 – 4.254	0.002
3+	4.321	2.707 – 6.898	<.001
Weekly hours listen radio			
Zero	---	----	----
1-19	1.738	1.158 – 2.609	0.014
20+	1.318	0.747 – 2.325	0.294
Weekly hours reading newspaper			
Zero	---	----	----
1-2	1.303	0.958 – 1.772	0.083
3+	1.700	1.214 – 2.380	0.007
Weekly hours on social networks			
Zero	---	----	----
1-20	1.096	0.831 – 1.445	0.468
21+	1.163	0.468 – 2.889	0.712
Weekly hours use internet			
Zero	---	----	----
1-20	1.149	0.905 – 1.459	0.217
21+	1.457	0.827 – 2.567	0.164

Notes: Categories used for the dependent variable defined by the response options given in that wave: 0 times, 1-2 times, 3-15 times, 16-30 times and 31+ times. N=1,986. p-values are for the hypothesis test that the odds ratio equals one.
Source: Wave 3 of CICPE survey.

Odds ratios greater than one indicate that the variable is associated with higher reported exposure levels. We see that residing in a treatment site was *not* significantly associated with higher reports – while the odds ratio is in the expected direction (1.l09), it is not statistically different from zero. Those aged 65 and older reported significantly lower levels of exposure than those aged 18 to 29, while those with at least some college reported significantly higher levels of campaign exposure than those without a high school degree. Additionally, the results indicate that in most cases more consumption of television, radio and print media also had significantly higher self-reported campaign exposure. However, the levels of social networking and internet use were not associated with self-reported campaign exposure reports.

That the model failed to establish a link between the heavy-up and self-reported campaign exposure is disappointing but does not necessarily mean the self reports are invalid. More likely, saturation played a part. Perhaps if the campaign had delivered lower overall control dosages, then the heavy-up signal would have been detected. Still, our findings do suggest that alternative measures and methodologies must be considered if an evaluation of the 2020 Census social marketing campaign is to take place. Very similar methodologies were used in both the Census 2000 and 2010 Census campaign evaluations – it would benefit the Census Bureau to consider very different approaches that rely less on self-reported exposure.

4.3 Ad Recall

A portion of the CICPE survey questionnaire was devoted to assessing the respondents' memory of and reaction to specific ads in the campaign. Respondents were screened early in the questionnaire by asking whether or not they had seen or heard anything about the Census recently. Those who answered positively were probed about four specific ads, a battery of overall reactions to all of the ads in the campaign, and their recall of specific ads using audio and visual prompts. These questions appeared in wave 3 only.

Because the DM campaign was targeted to all English speaking residents, it is likely that most members of the PAHUE sample were exposed to one or more of the DM ads. Consequently, all PAHUE respondents were asked about the first ad, titled "Frank," which was the single ad used most heavily during the Motivation phase of the DM campaign.

All participants in the PAHUE sample were also questioned about the second ad, "Miss Maybelle," without regard to their race or ethnicity. "Miss Maybelle" was created for use in the BAP.

The third ad, titled "Doors," was used in the NRFU phase of the campaign. The increase in media exposure for the PAHUE occurred during the Awareness and Motivation phases of the campaign. Heavy-up buys were not made during NRFU, so media exposure in treatment and control DMAs was more equal during that period. Consequently, we did not analyze the ad recall data for "Doors."

The fourth set of questions probed about an ad that was not used during the campaign—a placebo ad. This is the only ad probed in both waves 1 and 3. However, the description used for the placebo ad changed between waves 1 and 3. The description in wave 1 was somewhat generic because the advertising was still in development. The description in wave 3 used an ad that had been scripted, but not produced, as a template. This description was more in keeping

with the format and level of detail provided for the real ads in the campaign. For more information on the design of the ad recall questions and selection of the specific advertisements, see Datta, et al., 2012.

For each ad, the respondent was prompted with a short, general description of the ad, referenced below as a "verbal prompt." The purpose was to trigger the respondent's memory of the ad while not providing details of the characters, visuals, or copy in the ad. For those claiming to have recently seen the ad based on the short description, a followup open-ended question probed for any relevant details which the respondent recalled. The claim to have seen the ad was then classified as "confirmed" if the recalled details were sufficiently accurate. If the respondent indicated that they had recently seen the ad, they were asked to estimate how often they had seen it in the past 90 days, asked whether they thought the ad had "grabbed their attention," and whether they thought the ad gave them good reasons to mail back their census form. The exact wording of the questions varies slightly for each ad being probed.

Below are the questionnaire items asked for the placebo ads in wave 1 and 3 and the DM ad "Frank" in the wave 3:

> Let us start with advertisements you might have heard or seen in the media.
>
> Wave 1 – Placebo Ad:
>
> 19A_19A. Have you recently seen an advertisement that showed various film and music celebrities filling out their Census form or going to mail it in?
>
> 1☐ YES
> 2 ☐ MAYBE, NOT SURE
> 3 ☐ NO ➔ [SKIP REST OF PLACEBO QUESTIONS]
> 99☐ REFUSED ➔ [SKIP REST OF PLACBO QUESTIONS]
>
> Wave 3 – Placebo Ad:
>
> 19A4. Have you recently seen an advertisement that showed everything frozen in time (construction sites, school yards, hospital rooms, and chambers of congress) until a man dropped his census form into a mail box?
>
> 1 ☐ YES
> 2 ☐ MAYBE, NOT SURE
> 3 ☐ NO ➔ [SKIP REST OF PLACEBO QUESTIONS]
> 99☐ REFUSED ➔ [SKIP REST OF PLACBO QUESTIONS
>
> Wave 3 – DM Ad:
>
> 19A1. Have you recently seen an advertisement about the 2010 Census that shows a man sitting at home in his bathrobe filling out his census form?
>
> 1 ☐ YES
> 2 ☐ MAYBE, NOT SURE
> 3 ☐ NO ➔ [SKIP REST OF FRANK QUESTIONS]
> 99☐ REFUSED ➔ [SKIP REST OF FRANK QUESTIONS]

Follow-up probes for wave 3 ads (DM "Frank" ad, shown as example):

19B1. What happens in the ad? For example, can you tell me anything in detail about the people in the ad?

(DO NOT READ RESPONSE CATEGORIES AND CODE ALL THAT ARE MENTIONED.)

1. A middle-aged man in bathrobe and slippers taking his census form to his mail box
2. Phrase: "I'm just one guy"
3. Phrase: "but when I fill out the census form I'm helping to build a better school for Peter and Jen"
4. Phrase: "helping to improve roads for Mr Grippo's carpool"
5. Phrase: "Helping to improve healthcare"
6. Woman in labor on journey being wheeled to ambulance
7. Man walks down the street speaking into a megaphone and people of all walks gather around him
8. Phrase: "My Census answers help our voices to be heard in Washington"
9. Phrase: "so that we get our fair share of funding"
10. Standing at his mailbox surrounded by neighbors
11. Phrase: "Census is ten questions in ten minutes"
12. Puts form in mailbox in front of his house and laughs
13. Crowd of people cheer
14. Streamers come down from air
15. Phrase: "Fill it out and mail it back"
16. Phrase: "Help make your town better"
17. SLOGAN: We can't move forward until you mail it back
18. SLOGAN: Census 2010: It's in our Hands

66 ☐OTHER, specify: _____

77 ☐ DON'T KNOW

99 ☐ REFUSED

19C1. During the **past 90 days**, how often have you seen this advertisement? Would you say...?

1 ☐never

2 ☐once or twice

3 ☐ 3 to 15 times

4 ☐ 16 to 30 times

5 ☐ 31 times or more

77☐ DON'T KNOW

99☐ REFUSED

19D1. Would you say the ad grabbed your attention?

1 ☐ YES

2 ☐ NO

77☐ DON'T KNOW
99☐ REFUSED

19E1. Would you say the ad gave you good reasons to mail back your census form?
1 ☐ YES
2 ☐ NO
77☐ DON'T KNOW
99☐ REFUSED

The results of these probes should not be taken as literal estimates of the size of the population exposed to the advertising. Advertising uses visual and auditory stimulation to gain attention and create memory. It is likely that many who had seen the ad did not recall it from the simple, verbal prompt. Asking respondents to recall one or more of a set list of specific details in the ad may also raise the bar for "confirmed" recall (for more on a discussion of confirmed recall in the CICPE, see Datta, Hepburn, Yan and Evans, 2011).

As mentioned above, a placebo ad was probed in both waves 1 and 3 of the survey. The results for the two waves are shown in Table 17. The claimed recall of this ad increased significantly between waves. As we noted earlier, the verbal prompt description used for the placebo ad changed between waves 1 and 3. This change in the placebo description may have contributed to the increase in recall, i.e., the description sounded more like a real ad in wave 3 than in wave 1. There was no significant difference between the treatment and control cells in either wave. These results give a baseline against which to compare the reported awareness of ads actually used in the campaign. Because we are interested in differences across racial groups in response to the Miss Maybelle ad, we first examine whether there are racial differences in response to the placebo ad as a baseline. Table 18 displays the recall of the placebo ad in wave 3 for the non-Hispanic White and Black groups. The Black respondents show a slightly higher recall of the placebo ad.

Table 17. Ad Recall for Placebo Using Verbal Prompt, Treatment Versus Control

	Wave 1 (percent yes)	**Wave 3 (percent yes)**
Treatment	3.5 (1.19)	12.6*** (1.25)
Control	3.6 (0.65)	11.7*** (0.52)

p-values: *<.10, **<.05, ***<.01
Base: Respondents who reported having seen/heard something about the Census recently. Results are weighted.
Source: Waves 1 and 3 of CICPE survey.

Table 18. Differences by Race in Ad Recall for Placebo Using Verbal Prompt

	Wave 3 (percent yes)
Black Respondents	13.7* (1.10)
Non-Hispanic White Respondents	11.2 (0.60)

p-values: *<.10, **<.05, ***<.01
Base: Respondents who reported having seen/heard something about the Census recently. Results are weighted.
Source: Wave 3 of CICPE survey.

Turning to the real ads, as shown in Table 19, the treatment cell has a significant increase in both the recall of "Frank" and in recall of relevant details from "Frank" when compared to the control cell. The level of recall in the treatment cell is also significantly larger than the level observed for the placebo ad in wave 3 (Table 20). Table 19 also demonstrates that significantly more households in the treatment cell found that "Frank" grabbed their attention and gave good reasons to mail back the census form than in the control cell. Respondents in the treatment cell also reported that they had seen "Frank" at higher frequencies that those in the control cell. These findings are consistent with higher exposure to advertising in treatment sites.

Table 19. Ad Recall Measures for "Frank" Comparing Treatment and Control

	Recalled seeing the ad (percent yes)	Recalled details from the ad (percent with at least one detail coded)	Grabbed attention (percent yes)	Gave good reasons to mail census form (percent yes)	Recalled seeing the ad 3+ times (percent yes)
Treatment	17.6 *** (1.76)	7.6 ** (1.79)	14.0 *** (1.23)	12.8 ** (1.32)	11.4 ** (1.44)
Control	11.4 (1.14)	2.5 (0.53)	7.9 (1.57)	8.1 (1.17)	6.6 (0.41)

p-values: *<.10, **<.05, ***<.01
Base: Respondents who reported having seen/heard something about the Census recently. Results are weighted.
Source: Wave 3 of CICPE survey.

Table 20. Comparing Ad Recall for the Placebo and "Frank" Ads in Wave 3 Using Verbal Prompt

	Placebo – "Frozen" (percent yes)	"Frank" (percent yes)
Treatment	12.6 (1.25)	17.6** (1.76)
Control	11.7 (0.52)	11.4 (1.14)

p-values: *<.10, **<.05, ***<.01
Base: Respondents who reported having seen/heard something about the Census recently.
Results are weighted.
Source: Wave 3 of CICPE survey.

The second specific ad we examine, "Miss Maybelle," was part of the BAP. All PAHUE DMAs were exposed to this ad through targeted national cable time and network programming bought as part of the plan. The BAP levels were not systematically varied in the PAHUE. While there are no direct measurements of the GRPs associated with this ad, it is likely that they are at similar levels across all the DMAs. Four of the DMA pairs in the PAHUE received increased radio advertising as part of this targeted audience plan. The levels of radio advertising were roughly comparable between the treatment and control sides of each pair, except the control DMAs Columbia and Toledo received more BAP radio than their associated treatment DMAs, Savannah and Flint-Saginaw-Bay City. However, this discrepancy is unlikely to have an effect on these measures because the radio advertising did not use a version of "Miss Maybelle."

All respondents in the PAHUE sample of the survey were asked about their recall of and opinions about the "Miss Maybelle" ad. For the total sample, the treatment cell showed no significant difference in any of the recall measures for "Miss Maybelle" when compared to the control cell (Table 21). Because exposure to "Miss Maybelle" was through national media properties and not intentionally varied between the treatment and control DMAs, the results above are expected and support that the media delivery of "Miss Maybelle" was roughly equivalent.

Table 21. Total Sample Ad Recall Measures for "Miss Maybelle" for Treatment and Control

	Recalled seeing the ad (percent yes)	Recalled details from the ad (percent with at least one detail coded)	Grabbed attention (percent yes)	Gave good reasons to mail census form (percent yes)	Recalled seeing the ad 3+ times (percent yes)
Treatment	21.7 (1.05)	8.4 (2.21)	14.9 (2.23)	15.4 (2.08)	9.5 (1.45)
Control	22.7 (2.31)	7.0 (1.03)	16.4 (1.57)	14.9 (2.13)	8.7 (1.65)

p-values: *<.10, **<.05, ***<.01
Base: Respondents who reported having seen/heard something about the Census recently. Results are weighted.
Source: Wave 3 of CICPE survey.

Because "Miss Maybelle" was part of the BAP, it was aired in programming properties which have high ratings among the Black audience. The media selectivity of this programming is likely to generate relatively higher ratings among Black audience members. Ratings for the Black population for DM programming were measured to be 31 percent higher than the average ratings across the whole of the DM population. So members of the Black audience likely had higher levels of average exposure to both "Miss Maybelle" and "Frank" than members of the non-Hispanic White audience. This makes it interesting to compare the relative ad recall and other measures for these two ads between the different racial and ethnic groups with sufficient sample in the survey. As shown in Table 22, there are strongly significant differences in the recall, details reported, opinions of the ad, and self-reported frequency between the two audience groups for "Miss Maybelle," with the levels of recall and other measures for the ad being higher in the Black audience than among the non-Hispanic White audience.

Table 22. Ad Recall Measures for "Miss Maybelle" by Audience, Combining Treatment and Control

	Recalled seeing the ad (percent yes)	Recalled details from the ad (percent with at least one detail coded)	Grabbed attention (percent yes)	Gave good reasons to mail census form (percent yes)	Recalled seeing the ad 3+ times (percent yes)
Black	35.0*** (3.06)	17.2*** (2.23)	31.3*** (3.07)	30.8*** (2.64)	16.6** (1.88)
Non-Hispanic White	16.0 (1.12)	3.9 (1.56)	8.2 (1.37)	7.6 (1.02)	5.7 (0.79)

p-values: *<.10, **<.05, ***<.01
Base: Respondents who report having seen/heard something about the Census recently. Results are weighted.
Source: Wave 3 of CICPE survey.

Table 23 contrasts the recall and opinion measures for "Frank" for the two audiences. In contrast to the results for "Miss Maybelle," the recall of "Frank" is about equal and slightly above the rate observed for the placebo ad. The non-Hispanic White audience gave specific details from the ads at rates higher than that observed for the Black audience. However, there is no significant difference observed for agreement with the statements about "attention grabbing" and "reasons to mail the form." The non-Hispanic White audience self-reported frequency of exposure to "Frank" is not significantly different from that of the Black audience. In summary, these two audiences have very similar recall and other opinions of "Frank" in spite of a difference in exposure.

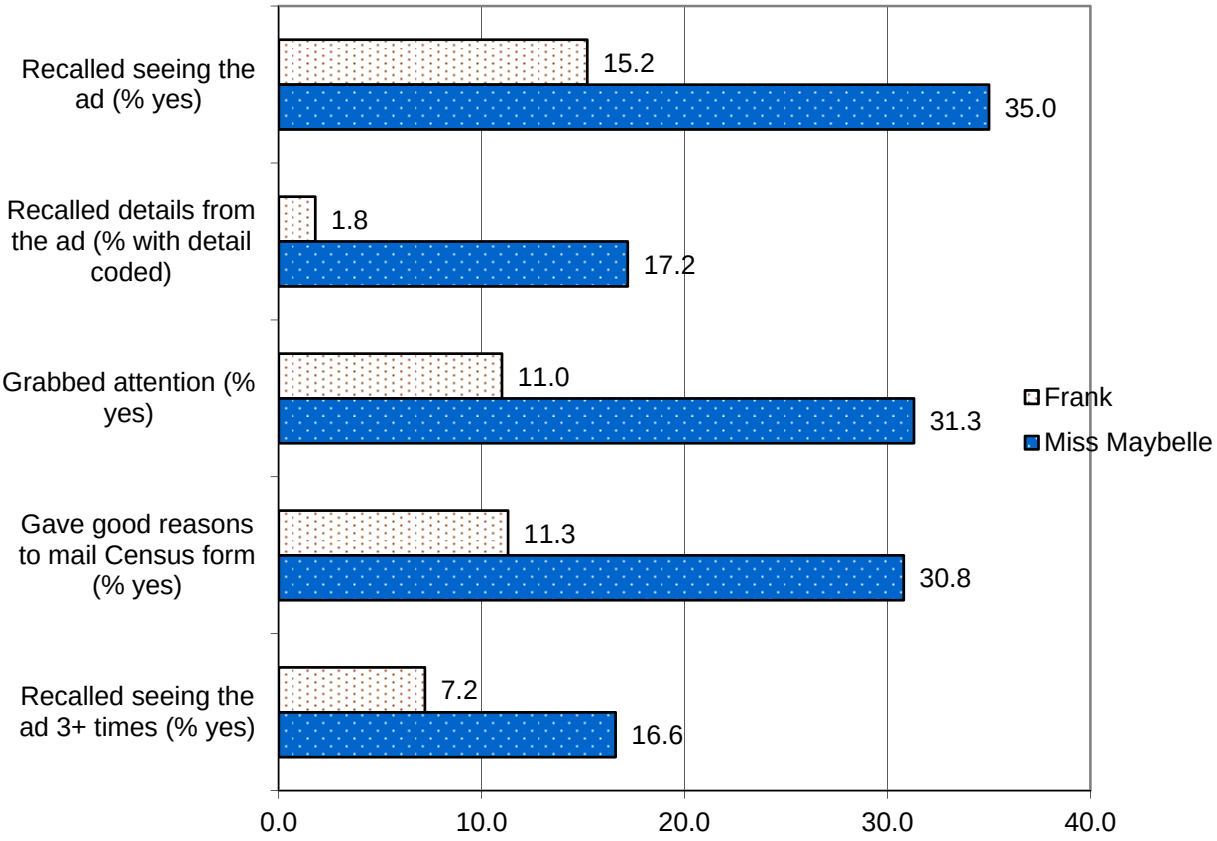
Table 23. Ad Recall Measures for "Frank" by Audience, Combining Treatment and Control

	Recalled seeing the ad (percent yes)	Recalled details from the ad (percent with at least one detail coded)	Grabbed attention (percent yes)	Gave good reasons to mail census form (percent yes)	Recalled seeing the ad 3+ times (percent yes)
Black	15.2 (3.19)	1.8 (0.99)	11.0 (1.76)	11.3 (2.39)	7.2 (1.26)
Non-Hispanic White	15.1 (1.97)	6.6** (1.92)	11.5 (2.42)	10.6 (1.87)	10.1 (1.60)

p-values: *<.10, **<.05, ***<.01
Base: Respondents who report having seen/heard something about the Census recently. Results are weighted.
Source: Wave 3 of CICPE survey.

Taken together, the findings in Tables 22 and 23, as displayed in Figure 2, point to an interesting possibility. "Miss Maybelle" appears to have higher levels of recall and other evaluative responses among the Black audience when compared to "Frank" in spite of the fact that the two ads had about equal exposure (GRPs) among the Black audience. There are many reasons this might occur – different relative effectiveness of the verbal prompts, differential creative quality of the ads, increased audience attention to advertising tied to the surrounding programming (i.e., media effects), etc. However, this finding should be considered in media planning for future campaigns.

Figure 2. Comparing Recall and Evaluation for "Miss Maybelle" and "Frank" Among the Black Audience



Base: Wave 3 Black audience respondents who report having seen/heard something about the Census recently. Results are weighted.

After prompting for individual ads, all respondents who reported having seen or heard something recently about the Census were asked a series of questions about their ratings of all of the ads and their assessment of the main topics and objectives of the overall campaign. Overall, there is only one significant difference in these measures between the treatment and control cells of the test, as shown in Table 24. Given that most treatment/control differences were not significant, it is unlikely that the extra advertising dosage in the treatment DMAs produced a strong incremental upward shift in agreement with these statements tied to the objectives of the campaign. It is unfortunate that these questions were not present in wave 1 so that we could investigate them for pre/post shifts. The figures shown below are for the percentage of participants responding in either of the top 2 boxes, "agree" or "strongly agree." However, treatment/control differences in the percentages of top box responses were also tested and none were found to be significant.

Table 24. Comparison of Statements about the Content of the Campaign – Percent Agree or Strongly Agree

Item	Treatment	Control
The Census 2010 ad campaign is about civic participation	88.8** (1.29)	83.2 (2.48)
The campaign is about gathering important information for our country.	94.1 (1.95)	92.8 (0.99)
The campaign is about helping people in my community.	90.0 (0.69)	88.5 (2.96)
The campaign is about making sure everyone like me is counted in the Census.	92.9 (1.67)	94.9 (0.94)
The campaign is about receiving our fair share of over $400 billion in federal funds for schools and other programs.	80.6 (2.54)	83.7 (3.28)
When I think of the Census 2010 ad campaign, I think returning the census form is the right thing to do.	95.9 (0.66)	94.3 (1.46)
When I think of the campaign, I think people in my community are all returning the census form.	48.8 (3.29)	52.4 (2.78)
When I think of the campaign, I think it's just 10 simple questions and takes about 10 minutes	90.9 (1.48)	89.3 (1.65)
When I think of the campaign, I think mailing back the form helps the country move forward.	87.4 (1.69)	90.9 (2.69)
When I think of the campaign, I think I can make a difference.	79.6 (2.44)	82.2 (3.10)
When I think of the campaign, I think returning the form helps communities get better healthcare, education and job training.	81.1 (0.99)	84.0 (3.35)

p-values: *<.10, **<.05, ***<.01
Base: Respondents who report having seen/heard a Census paid advertisement on at least one medium. Results are weighted.
Source: Wave 3 of CICPE survey.

Wave 3 of the CICPE contained a section which probed how audio or visual stimulation might improve recall of the ads when compared to a verbal prompt. Using video and/or visual story board prompts is a common industry practice in research for advertising tracking. This section was added to the survey after the other sections in the questionnaire concerned with recall of ads using verbal prompts (reported above), partnership communications, Census in the Schools, and communication via earned media. The audio/visual prompts were drawn from the "Frank" ad. Interviews completed via telephone used an extract from the ad's sound track as an audio

prompt. It was edited to remove references to detail which could be used for assessing actual exposure to the ad. Interviews completed in person or on the web used a story board of still images taken from the ad.

There are many issues with this "test" of advertising stimuli. The two modes of stimulation, audio and still visuals, are confounded with CATI (audio) and paper and pencil interview PAPI (stills) interviewing. The interview modes may be correlated with the difficulty in gaining respondent cooperation. There was no a priori experimental estimation of the relative effectiveness of the two modes of stimulation among respondents known to have been exposed to the ad. The verbal prompt earlier in the survey may have enhanced responsiveness to the audio/visual prompts. We recognize these factors when looking at the survey results.

Both the audio and still stimuli produce higher reported recalls than the verbal prompt used earlier in the questionnaire. This is in line with industry experience concerning the desirability of using more content-rich stimuli when cueing ad recall. The treatment DMAs show a higher level of recall for both the audio and verbal stimuli, as shown in Table 25. Interestingly, there is no significant difference between treatment and control DMAs for the still visual stimulus. Assuming that the stimuli equally represent the ad, this could indicate that the sound track in "Frank" was more memorable than the visuals. It is less likely that difference in recall is related to the interview method, CATI or PAPI, or the respondents interviewed by either method. The treatment cells are significantly different from the control cells for either method when using the verbal prompt.

Table 25. Recall of "Frank" Using an Audio or Visual Stimulus -- Percent Responding Yes

Stimulus	Treatment	Control
Stills	36.1 (4.24)	31.4 (2.77)
Audio	48.3** (2.57)	36.9 (2.76)
Verbal prompt (Q19A1)	17.6 *** (1.76)	11.4 (1.14)
Verbal prompt – CATI (Q19A1)	28.8 * (2.70)	16.5 (4.05)
Verbal prompt – PAPI (Q19A1)	16.1 ** (1.71)	10.7 (1.52)

p-values: *<.10, **<.05, ***<.01
Base: Respondents who report having seen/heard something about the Census recently. Results are weighted.
Source: Wave 3 of CICPE survey.

The survey probed whether those who recalled the ad via the audio or visual prompt found it convincing or not. There was no significant difference between the treatment and control DMAs on this measure, as shown in Table 26.

Table 26. Comparison of Agreement that "Frank" Is Convincing Using an Audio or Visual Stimulus -- Percent Responding Yes

Stimulus	Treatment	Control
Stills	69.2 (6.05)	61.3 (5.28)
Audio	79.4 (3.93)	83.2 (2.27)

p-values: *<.10, **<.05, ***<.01
Base: Respondents who reported having seen "Frank" after exposure to either the audio or visual prompt.
Results are weighted.
Source: Wave 3 of CICPE survey.

Going forward, the Census Bureau should strongly consider use of audio and visual stimuli for ad tracking and surveys. This approach would require advance coordination with the Office of Management Budget when obtaining research approvals because the actual ad stimuli might not be available well in advance of the start of interviewing. The tight intervals between final ad development and on-air dates are a natural consequence of the advertising development and approval process. There is a strong desire to ensure that the final advertising is current in its topical approach, uses up-to-date creative methods, reflects management input, and utilizes the results of pre-testing research. That said, tracking research and its associated learning would benefit from a more reliable method to identify respondents who had seen the advertising.

4.4 Attitudes About the Census

In this section, we explore the degree to which attitudes and knowledge about the census may have shifted over the course of the campaign. Of the battery of questions examined here, some are very census-specific and run the gamut from knowledge of what census data are used for, to personal opinion on the census' importance, to degree of trust regarding data confidentiality. A few are arguably more aligned with core beliefs and values (such as level of trust in government) while others are not very salient to the general public and may come closer to what Bishop (1980) and others describe as "nonattitudes" or "pseudo-opinions." The former may be more strongly held and resistant to change as a result of a communication campaign. For example, throughout the 2010 Census ICC, the Census Bureau asked ten questions on a daily telephone tracking survey. Among the questions were items asking whether it is important to count everyone in the census and whether the Census Bureau's promise of confidentiality can be trusted. For each, opinions shifted little from December 2009 to late April 2010 – the entire span of the communication campaign (Miller and Walejko, 2010).

The PAHUE involved ads that were part of the DM plan – it did not increase rotation of ads in the Black, Hispanic, in-language or other targeted audience plans. Consequently, the variation in the heavy-up creative catalog was limited. In all, twelve television ads were increased in rotation as part of the heavy-up. Three ads were part of a sequential series directed by Christopher Guest. These were designed to raise awareness (as opposed to a call to action). Accordingly, they

mostly ran early in the campaign during the awareness phase. The slogan was "A snapshot of America."

The remaining TV ads during the awareness and motivation phases fell mostly into two groups: "Frank" (motivation phase only) and "X's and O's" (both awareness and motivation phases). The two were visually very different from one another but both contained the same two core messages: (1) the census has only ten questions and takes only ten minutes and (2) completing the form will help your community get its fair share of federal funding for things like hospitals, school and roads. Featured in the ads were phrases such as "so that we can get our fair share of funding," "census is ten questions in ten minutes," "if we don't know how many people there are, how do we know how many roads we need?" and "we can't move forward until you mail it back." The main messages for the heavy-up radio, out-of-home, digital and print ads were the same -- that the census was quick and easy and critical for communities to know their needs. Typical phrases from radio and print ads were "It's 10 simple questions and takes about ten minutes," "it's how we get out fair share of funding for the things we need," "if we don't know how many people we have, how do we know how many carpool lanes we need?" and "we can't move forward until you mail it back." Consequently, while each ad had its own look and feel, the messaging was consistent across different ads and media.

The CICPE survey measured knowledge and attitudes by a series of yes/no items about census uses, a battery of opinion statements about the census, and a question whether census participation is required by law. These questions were entered into an exploratory factor analysis as a way to reduce the number of items into a smaller number of underlying constructs.[9] The analysis revealed four major factors which together explained about 48 percent of the variance across the knowledge and attitude items. Factor 1, which we labeled "Privacy /Confidentiality Concerns" had high loadings on four items. Factor 2 labeled "Census is Necessary" had high loadings with four questions, and a third factor labeled "Tracking/policing activities" loaded high with three questions. The final factor labeled "Allocate resources" had high loadings on three items. See below for description of the items loading highest on each factor:

(F1) Privacy / Confidentiality Concerns:
Q 16B (Census is invasion of privacy)
Q16C (Promise of confidentiality can be trusted) – RECODED TO NEGATIVE
Q16D (Information will be misused)
Q16F (Answers could be used against me).

(F2) Census is necessary / important / not burdensome:
Q16G (Census matters for my family)
Q16H (Government already has my information) – RECODED TO POSITIVE
Q16I (Doesn't matter to me personally) – RECODED TO POSITIVE
Q16J (Takes too long to fill out) – RECODED TO POSITIVE.

[9] Because the proportion of respondents answering "don't know" was nontrivial for some of the Census use items (Q15A-G), answers of "don't know" were recoded to the "incorrect" category and retained in the factor analysis. For the opinion items (Q16A-J), answers of "neither agree nor disagree" were recoded to the midpoint of the five point agree/disagree scale – "neither" was not explicitly offered as a response option but was accepted if volunteered.

(F3) Tracking / policing:

Q15D (Used to determine property taxes)
Q15E (Used to find lawbreakers)
Q15G (Used to find illegal residents).

(F4) Allocate resources:

Q15A (Used to decide money for communities)
Q15B (Used to decide representatives in Congress)
Q15F (Used to help businesses and governments plan for future).

We created two sets of scores for the four factors: standardized factor scoring coefficients with loadings on *all* of the variables included in the factor analysis and factor indices, using only the items with high loadings on the corresponding factor. In the latter case, the score was simply the sum of the items that made up the index and items were recoded to run in the same direction (either positive or negative). For example, for Factor 1, Q16C was recoded such that a high score reflected *mistrust* of the Census Bureau's promise of confidentiality. Results based on the factor scores and factor indices yielded the same conclusions and the correlations between the two were very high (ranging from .89 to .96). As a result, and for simplicity, we opted to use the mean factor indices for our analysis and presentation here.

Given the type of messages contained in the DM base ads, we hypothesized there might be a detectable change between wave 1 and wave 3 on the Factor 2 and Factor 4 indices as these came closest to representing the ad themes. Given the heavy-up messaging, we had little reason to expect differences in the privacy or policing attitude indices.

Table 27. Mean Scores to Census Attitude Indices, Treatment Versus Control

	(F1) Privacy/Confidentiality Concerns (index mean)		**(F2) Census Necessary / Important** (index mean)		**(F3) Census Used for Tracking / Policing** (index mean)		**(F4) Census Used to Allocate Resources** (index mean)	
	W1	W3	W1	W3	W1	W3	W1	W3
Treatment (T)	9.1	8.8	14.9	15.7	4.3	4.4	5.2	5.3
	(.04)	(.15)	(.07)	(.17)	(.04)	(.06)	(.02)	(.04)
Control (C)	9.1	9.0	15.2	15.6	4.4	4.4	5.2	5.3
	(.12)	(.15)	(.10)	(.14)	(.05)	(.06)	(.04)	(.04)
TEST	**Estimate**		**Estimate**		**Estimate**		**Estimate**	
W1(T) – W1 (C)	.10		-.29		-.09		-.03	
	(.13)		(.16)		(.07)		(.04)	
W3 minus W1 (C)	-.06		.46**		-.06		.12	
	(.22)		(.14)		(.06)		(.07)	
Change in (T) minus change in (C)	-.33		.38		.18		-.02	
	(.19)		(.30)		(.10)		(.09)	

p-values: *<.10, **<.05, ***<.01
Base: Respondents reporting ever heard of Census. Results are weighted.
Source: Waves 1 and 3 of CICPE survey.

Table 27 contains the four factor index means at wave 1 and wave 3 for the treatment and control. On all four attitude indices, the treatment and control means started at similar points with no detectable differences at wave 1. For the privacy/confidentiality index, we saw no significant decrease in confidentiality concerns between waves for the control group. Additionally, the difference in differences between the wave 1 and wave 3 treatment and controls was also just nonsignificant (p=.12). As hypothesized, we found no evidence the DM messaging did anything to reduce concerns around data misuse, privacy and confidentiality.

On Factor 2, the control cases saw a significant increase between wave 1 and wave 3 in the perception that the census is necessary, matters, and quick to complete. Again, however, while the difference between the treatment and control differences was in the expected direction (the treatment saw a larger absolute increase in this perception), it was not large enough to be statistically significant (p=.24).

For the policing/tracking index (Factor 3), there was no significant shift between waves for the control group nor was the wave 1-wave 3 difference between treatment and control statistically significant. This was not unexpected given the heavy-up messaging did not attempt to debunk any of the misperceptions around census uses. Tests for Factor 4 (Allocate resources) yielded similar results – no changes between wave 1 and wave 3. In areas receiving the normal dosage of advertising, the belief that census data are used to allocate community resources and decide needs of the community did not significantly increase over the course of the ad campaign. Likewise the degree of change between treatment and control over the campaign was also not significant. The lack of change over time for this factor is somewhat puzzling given that messages in the television, radio, and print ads certainly stressed the connection between the census and delivery of community services. From the Census Continuous Tracking Survey (CCTS), we know the percentage that reported hearing that "information from census forms is used to decide how $400 billion in federal money is spent on schools and other programs" increased from around 25 percent to 60 percent between early December and the third week of April (Miller and Walejko, 2011).

Knowing this, we examined individually the three items making up this index. We found that neither CONGRESS (Q15B) nor PLAN (Q15F) awareness changed between waves but that the percentage believing the census is used to decide how much money communities receive from the government (MONEY Q15A) did increase significantly (from 63.5 percent to 80.2 percent in the treatment and from 65.5 percent to 79.3 percent in the control). This suggests that overall, the campaign was successful in driving this theme home, but again, the extra dosage of advertising did not make any difference.

A few of the opinion/knowledge items did not load highly with any of the four factors. These included Q12 (general feelings about Census), Q13 (familiar with how census impacts community), Q14 (does law require census participation), Q15c (taking part shows pride), and Q16e (census counts both citizens and non-citizens). We examined these individually and detected no differences between waves with one exception. In both the treatment and control sites, the percent of respondents who (correctly) indicated that census participation is required by law increased significantly from wave 1 to wave 3. At wave 1 in both sites, just over 20 percent said census is required by law, but by wave 3, over half of the respondents in both sites knew the

census was mandatory (data not shown)[10]. The increase between waves was not different between the treatment and control but again, this was not unexpected given the messaging. More likely this message was communicated by way of the census form itself – the message "Your response required by law" was prominently displayed on the front of the envelope.

4.5 Mail Return Behavior

In wave 3, respondents were asked if they received a form in the mail and if yes, whether they had completed and mailed it back or not. Additionally, following the close of the wave 3 data collection, the Census Bureau supplied NORC with a file containing information on how the heavy-up sample households responded to the 2010 Census (e.g., by mail, personal enumeration, or by some other means). From this, NORC created a census mailback variable that corresponded to actual behavior. For a more detailed methodology, see (Datta et al., 2012).

In this section, we concentrate on actual mail back behavior. In particular, we focus on whether households in areas slated to receive a mail form mailed it back *before* the cut for personal nonresponse followup (NRFU). We felt this was the most appropriate outcome measure because we are primarily interested in understanding the relationship between exposure to paid advertising and returning a form in time to avoid to sending an in-person enumerator (which is far more costly than mail). However, we offer the caveat that while this measure is arguably better than self-reported behavior, our outcomes only reflect households that agreed to participate in the CICPE surveys. We feel that the best analysis of mailback behavior is contained in a later section of this report (section 4.6). There, we will present a mail return rate analysis from all households contained within treatment and control heavy-up DMA sites, not just those that participated in the CICPE surveys. Here, we examine differences in CICPE survey participant's mailback rates between the heavy-up treatment and control groups and by race and ethnicity and audience segments within the treatment and control. We also look at self-reported exposure level by mail back behavior within the treatment and control.

Table 28. Percent Mailing Back Census Form Prior to NRFU, Treatment Versus Control

	Percent Mailed Form Before NRFU
Treatment (T)	66.7
	(2.3)
Control (C)	64.0
	(2.0)

p-values: *<.10, **<.05, ***<.01
Base: Pooled heavy-up cases eligible to receive a mail form that were geocoded to Census operational file containing mail return information. Weighted results.
Source: Waves 1 and 3 of CICPE survey matched to Census operational data on mail return status of housing unit.

[10] For detailed results of individual questionnaire items between wave 1 and wave 3, see Datta, Hepburn, Yan and Evans, 2011.

Table 28 presents the percent who returned a census form prior to the personal-visit NRFU operation. For this table we present both the wave 1 and wave 3 respondents pooled – there is no compelling reason to examine separately by waves since both received the heavy-up dosage in the same manner regardless of the interview wave they participated in. Further, the wave 3 data collection did not begin until after the start of NRFU and therefore could not have influenced a household's likelihood to mail back a form prior to that date.

We see that the difference in mailback rates is in the expected direction with the treatment having a higher level (66.7 percent) compared to the control (64.0 percent), however, the difference between groups is not significantly different from zero. We look at mail return behavior across the treatment and control sites by race and ethnicity (Table 29) and audience segmentation cluster (Table 30). This seeks to understand if the relationship between increased frequency of advertising and mailback behavior might be conditional upon certain demographic and socioeconomic characteristics. This hypothesis is explored more rigorously in a later section of the report (Section 4.6.3), but we present some of the preliminary survey-based findings here.

Table 29. Percent Mailing Back Form Prior to NRFU, Treatment Versus Control by Race/Ethnicity

	Treatment (Percent mailed form prior to NRFU)	Control (Percent mailed form prior to NRFU)
White/Other race, non-Hispanic	70.8	68.5
	(2.5)	(1.9)
Black, non-Hispanic	58.8	57.3
	(3.0)	(1.4)
Hispanic, any race	55.1	45.9
	(5.8)	(4.4)

p-values: *<.10, **<.05, ***<.01

Base: Pooled heavy-up cases eligible to receive a mail form that were geocoded to Census operational file containing mail return information.

Source: Waves 1 and 3 of CICPE survey matched to Census operational data on mail return status of housing unit.

While the difference in mailback of forms was in the expected direction for all race and ethnic groups, none of the subgroup differences were statistically significant between the treatment and control. Table 30 presents mailback rates across the audience clusters used to plan and target the 2010 communication campaign. The Census Bureau developed these segments using Census 2000 tract-level demographic and housing characteristics, Census 2000 mail return behavior, information from the American Community Survey, and consumer survey data (see Bates and Mulry, 2011). Each household responding in the CICPE was mapped back to its audience segment based on address. In all, eight segments were produced for purposes of planning the communications campaign, but several had to be combined and two excluded in the analysis due to heavy-up sample size constraints. [11]

[11] The All Around Average (owner skewed) and All Around Average (renter skewed) were combined as were categories of Economically Disadvantaged (owner skewed) and Economically Disadvantaged (renter skewed). Both Ethnic Enclave (owner and renter skewed) and Single Unattached Mobile categories were excluded because the heavy-up sample did not have large enough representation in these tracts.

Table 30. Percent Mailing Back Form Prior to NRFU, Treatment Versus Control by Audience Segment

	Treatment (Percent mailed form prior to NRFU)	Control (Percent mailed form prior to NRFU)
Advantaged Homeowners	76.4	68.9*
	(2.6)	(2.3)
All Around Average	63.7	66.2
(owner+renter skew)	(1.9)	(2.0)
Economically Disadvantaged	66.3	54.9***
(owner+renter skew)	(4.3)	(1.9)

p-values*<.10, **<.05, ***<.01

Base: Pooled heavy-up cases eligible to receive a mail form that were successfully geocoded to Census operational file containing mail return information.

Source: Waves 1 and 3 of CICPE survey matched to Census operational data on mail return status of housing unit.

For two of the segments (Advantaged Homeowners and Economically Disadvantaged), the mailback rate was higher in the treatment areas compared to the control. These differences ranged from 7 to 11 percent and represent the only differences found in this section to be statistically different between the treatment and control. This is the first evidence suggesting the extra advertising may have influenced behavior. Because the audience segments are composites of 12 variables making up a "hard-to-count" score, they are more robust indicators than any single characteristic. It is noteworthy that the two segments where differences emerged are fundamentally different – the Advantaged Homeowner (AH) segment is characterized by homeowners with low mobility living in primarily single-unit, spousal-occupied households. Conversely, the Economically Disadvantaged (ED) segment is skewed toward non-spousal households living in poverty where unemployment is above average, the presence of children is above average, and the units are typically multi-units located in urban areas.

For the AH segment, the extra dosage likely paid dividends because the experiment doubled advertising in the DM plan which includes anyone who consumes English language media. Further, the AH group has a historically high response rate and is inclined to respond but the frequent advertising probably served as a reminder and helped overcome any lingering inertia. This treatment effect is particularly noteworthy because this segment is one of the largest, comprising over one-quarter of all U.S. households.

The additional advertising effect among the ED segment is perhaps more complex, involving other components of the campaign. This group was heavily targeted in both the Black and DM advertising plans and was also targeted by the community-based Partnership Program. Further, the ED group is also a large consumer of cable television (which comprised the majority of the heavy-up catalog). Each of these elements taken together, along with the extra dosage of paid advertising, may have interacted in a way to increase mail response.

Table 31. Percent Mailing Back Form Prior to NRFU, Treatment Versus Control by Level of Exposure to Paid Ads and Earned Media

	Treatment (Percent mailed form prior to NRFU)	Control (Percent mailed form prior to NRFU)
Low exposure	65.9 (3.0)	61.6 (3.8)
High exposure	69.9 (1.9)	66.9 (1.8)

p-values: *<.10, **<.05, ***<.01

Base: Heavy-up cases eligible to receive a mail form that were geocoded to Census operational file containing mail return information. Those who skipped the campaign exposure frequency questions assigned to the 'low exposure" category.

Source: Wave 3 of CICPE survey matched to Census operational data on mail return status of housing unit.

In Table 31, we present mailback rates by high and low self-reported campaign exposure at wave 3 (reminding readers of the earlier discussion around self-reported advertising recall and measurement error). Campaign exposure here is defined as a combination of exposure to paid advertisement and features in earned media such as newspaper articles and television news. The high exposure category includes those who saw or heard a paid ad three or more times and saw one or more census news stories within the last three months – the low exposure is anything below this threshold. Recall that earlier we reported finding no significant difference in the wave 1 to wave 3 change in frequency of self-reported campaign exposure between the treatment and control sites. The story is similar here, that is, while the mail return percents are higher in the treatment for both exposure categories, neither difference is statistically different from zero. This lack of difference could be an outgrowth of the fact that self-reported exposure and actual exposure are often quite different. Survey questions are a less-than-perfect way to delineate low versus high exposure and our measure may lack the sensitivity to adequately pick up dosage differences. However, despite shortcomings of the measure, the *direction* of the relationship between exposure level and mailback behavior within the treatment and control is as expected – that is, higher absolute mailback rates in both categories in the treatment compared to the control (although again, the difference is not statistically different from zero).

4.6 Analysis of Census Operational Data

In this section, we analyze data from all households within the DMAs involved in the PAHUE. This analysis is based on the 4.96 million households in these DMAs who were eligible to return a form by mail. We exclude housing units that were determined to be vacant on April 1st, and housing units for which census received a returned form from the post office labeled "undeliverable as addressed" (UAA). We focus on mail return rates (MRR)—the share of occupied housing units that returned a form by some specified time—and use April 19th as the cutoff date. This date was chosen because any address that had a completed form returned by that date would not be sent to NRFU.

4.6.1 Mean Differences in Mail Return Rates

The combined mail return rate across all control sites was 75.3 percent and was 74.9 percent in the treatment sites. Table 32 presents the Census mail return rates for each DMA pair. Results were mixed. For three of the pairs, the rates are in the expected direction with the treatment sites

having higher mail returns. For another three pairs, the control sites had higher returns and for two pairs (Tallahassee-Shreveport and Little Rock-Jacksonville), the difference was negligible (less then one-half of a percent). It is interesting to note that the pair we believe came closest to executing the experiment as originally planned (Joplin-Erie), had higher mail returns in the treatment site.

Table 32. Mail Return Rates for the Eight PAHUE DMA Pairs

Control Site	Mail Return Rate	Treatment Site	Mail Return Rate	Treatment-Control Difference
1. Columbia	78.3%	Savannah	74.4%	-3.9
2. Toledo	78.6%	Flint-Saginaw	83.4%	4.8
3. Jackson	72.4%	Montgomery	73.1%	0.7
4. Augusta	76.0%	Baton Rouge	69.2%	-6.8
5. Tallahassee	72.6%	Shreveport	72.4%	-0.2
6. Lubbock	74.9%	Odessa	72.0%	-2.9
7. Joplin	76.3%	Erie	82.0%	5.7
8. Little Rock	73.1%	Jacksonville	72.8%	-0.3

Source: 2010 Mail Response/Return Assessment data for PAHUE areas.

Table 33. Mail Return Rates by Audience Segment, Treatment Versus Control

Audience Segmentation Cluster	Control MRR	Treatment MRR	Treatment-Control Difference
All Around Average (homeowner skew)	75.9%	75.1%	-0.8
All Around Average (renter skew)	72.0%	71.2%	-0.8
Economically Disadvantaged (homeowner skew)	70.5%	69.3%	-1.2
Economically Disadvantaged (renter skew)	63.0%	64.8%	1.8
Ethnic Enclave (homeowner skew)	72.2%	70.2%	-2.0
Ethnic Enclave (renter skew)*	n/a	n/a	n/a
Mobile singles	65.8%	66.8%	1.0
Advantaged Homeowners	81.2%	81.6%	0.4

* Too few tracts in the PAHUE fell into this cluster to be represented here.
Source: 2010 Mail Response/Return Assessment data for PAHUE areas.

Table 33 presents the mail return rates by audience segments by treatment and control. The difference in mail return rates between treatment and control are small for most segments – usually less than one percent. The Economically Disadvantaged renter skewed households had slightly higher mail returns in the treatment areas while the Ethnic Enclave homeowner skewed had slightly lower.

One of the limitations of the PAHUE and most in-market media tests is the small number of test pairs. As such, the analysis is sensitive to DMA-to-DMA differences that would exist in the absence of any experimental treatment. Some across-DMA variation in return rates may be due to factors we can measure—for example, the size of the foreign-born population or activities and outreach of the local 2010 Partnership Program. If large differences exist, it can reduce the ability to detect significant difference in mail return associated with the treatment. One way to explore this is to examine difference in MRRs between treatment and control DMAs for narrower groups.

Pursuing such a strategy, we examined mail return rates for each DMA by race/ethnic group and by audience segmentation clusters (see Appendix C). The variation in mail return rates between the control sites is large for some groups. For example, mail return for the Black population ranged from a low of 61.3 percent in Little Rock to as high as 73.8 percent in Columbia. Likewise, for tracts representing the Economically Disadvantaged homeowner skewed areas, mail return rates from the control sites ranged from as low as 68.2 percent (Jackson) to as high as 75.5 percent (Columbia). This suggests that other covariates are probably in play and adding noise to our ability to detect the heavy-up's impact on differences in mail return associated with the treatment dosages. But with many potentially inter-related factors, a strategy of examining one or two factors at a time is unlikely to make clear whether the differences we know exist in these areas are confounding the pattern we expected to find. With this in mind, in the next section we use multivariate statistical models to control for an extensive set of characteristics that may influence the likelihood of a mail return.

4.6.2 Logistic Models of Mail Returns

Our goal in this section is to determine if there is a significant difference in response rates across treatment and control sites when we compare households with similar characteristics. We follow the same basic set of comparisons used above, in which we first present results based on all households. We then examine whether the results differ if we divide households into groups defined by the race and ethnicity of the householder. And finally we examine whether the results vary across groups defined by the audience segmentation cluster of the tract in which they reside.

We carry out this analysis using logistic models, in which the dependent variable equals 1 if the household returned a form by April 19, 2010, and 0 if not. In the results presented here, we use the logarithm of the GRPs associated with a householder's race/ethnic group in the DMA they live in as our measure of the experimental treatment. We have also estimated models that parallel those presented here but use (i) an indicator variable for treatment sites, and (ii) use GRPs without taking the logarithm. The three alternatives we investigated gave quite similar results, and we chose to present the log GRP results because the assumption underlying this functional form is more consistent with the idea that there are decreasing returns to increasing the level of advertising. The log GRP form implies, for example, that increasing GRPs by 100 points has a much larger effect on mail return rates when moving from 500 to 600 GRPs than it does when moving from 5000 to 5100 GRPs.

We include indicator variables for the relevant DMA pair for each household in order to account for mean differences in response rates across pairs. This leaves only within-pair MRR differences to be picked up by the log GRP variable. We include a large number of household

characteristics, such as whether the household rents or owns its residence, whether it is headed by a married couple, householder age, householder race/ethnicity, the number of household residents, and how many of them are children. We also include several controls for differences in how an address was treated during survey operations: whether the household was scheduled to receive a replacement form, whether it received a bilingual form, and how forms were delivered to the address.[12]

Finally, we include a measure of the level of partnership activity in the county in which each household lives. Census partnered with a wide range of groups to increase awareness of the census and encourage households to return their forms. While precisely measuring differences across areas in the levels of this activity would be very difficult, the Census Bureau did maintain a database that included county-level information on activity such as the number of active partners, the dollar value of grants awarded for partnership support, and counts of local partnership staff. Here we used a scale measure based on these data that was developed by NORC in consultation with the Census Bureau's Field Division (see Datta, et al., 2012 for details). It is constructed so that a one unit increase in the scale represents a one standard deviation increase in the level of measured partnership activity in the household's county of residence.

[12] Note, in areas in which replacement forms were sent only to non-responding addresses, we control for eligibility for a replacement form, not whether one was sent. Using this approach is to be preferred because an indicator that a replacement was sent would be 0 for all households that returned a form early in the mail-back period, and all households that did not end up returning a form pre-NRFU would have a value of 1. Thus the indicator would be likely to have a negative coefficient even when replacement forms had a strong positive effect on the overall response rate simply because it would identify the households that are less likely to respond.

Table 34. Logistic Model of Mail Return Rate, All Eligible Households

Dependent var = Valid mail return by 4/19/10	Odds Ratio		95% Confidence Interval
Log (GRP)	1.013		[0.957,1.072]
Type of enumeration area			
Update/leave enumeration area	0.710	***	[0.644,0.783]
Military enumeration area	1.035		[0.869,1.233]
Urban update/leave enumeration area	0.619	***	[0.554,0.691]
Outcome of 2000 collection at this address			
Address not in 2000 census	0.824	***	[0.800,0.849]
Enumerator completed return in 2000	0.635	***	[0.623,0.648]
Other outcome in 2000	0.719	***	[0.705,0.734]
Type of housing unit			
Multi-unit building	0.857	***	[0.836,0.878]
Trailer/mobile home/boat/RV, etc	0.737	***	[0.714,0.761]
Married couple	1.608	***	[1.587,1.630]
Household rents or occupies without paying rent	0.627	***	[0.613,0.641]
DMA pair for PAHUE			
Savannah/Columbia SC	1.244	*	[1.010,1.534]
Flint-Saginaw/Toledo	1.433	*	[1.044,1.967]
Selma/Jackson MS	1.043		[0.976,1.114]
Baton Rouge/Augusta	0.998		[0.832,1.198]
Shreveport/Tallahassee	1.098	*	[1.005,1.199]
Odessa-Midland/Lubbock	1.192		[0.858,1.654]
Erie/Joplin-Pittsburg	1.298		[0.992,1.700]
Household of unrelated people	0.717	***	[0.688,0.747]
Partnership activity scale (#1)	0.999		[0.991,1.007]
Age of householder			
Householder age <=25	0.744	***	[0.720,0.769]
Householder age 41-64	1.596	***	[1.561,1.632]
Householder age 65+	2.991	***	[2.847,3.141]
Number of children age<18 in household			
1 child	0.959	**	[0.940,0.979]
2 children	1.008		[0.994,1.022]
3-4 children	0.978		[0.956,1.000]
5-8 children	0.870	***	[0.826,0.917]
9+ children	0.745		[0.470,1.182]
Number of residents in household			
2 people	1.181	***	[1.171,1.191]
3 people	1.033	**	[1.014,1.052]
4 people	0.939	***	[0.921,0.958]
5 people	0.820	***	[0.801,0.839]

Table 34. Logistic Model of Mail Return Rate, All Eligible Households (continued)

Dependent var = Valid mail return by 4/19/10	Odds Ratio		95% Confidence Interval
9-12 people	0.885	**	[0.827,0.947]
13+ people	0.31	***	[0.222,0.434]
In area blanketed by replacement forms	1.215	***	[1.146,1.289]
In area where non-respondents received replacement form	1.128	***	[1.075,1.184]
Address is not residential	0.035	***	[0.026,0.047]
Non-Hispanic Black	0.748	***	[0.690,0.811]
Hispanic	0.607	***	[0.526,0.700]
Received bilingual form (English/Spanish)	1.045		[0.957,1.141]

Note: Omitted categories: Mail-out/mail-back enumeration area, address had respondent return in 2000, owner, single-family house, Jacksonville/Little Rock DMA pair, single person household, non-Hispanic White, householder aged 26-40.

p-values: *p<0.05, **p<0.01, ***p<0.001
Source: 2010 Mail Response/Return Assessment data for PAHUE areas.

Table 34 presents odds ratios for the effect of each of the variables included as controls in our model. The first line gives the odds ratio and confidence interval around it for log GRP. As was true with the ordered logistic results presented earlier, an odds ratio above 1 means that a characteristic is associated with a higher value of the dependent variable, or in this case, with a higher likelihood that a household returns a form prior to NRFU. If the confidence interval around the odds ratio includes 1 in its range, then the effect of that control is not significant. Here the odds ratio on the log GRP is slightly above 1, but is not significantly different from 1.

A number of characteristics are associated with significant differences. For example, we include controls for the form of response from the same housing unit in 2000, with the omitted category being housing units with a 2000 mail return. Each of the other possibilities—the address did not exist in 2000, an enumerator completed return, or some other non-mail-return outcome—is associated with a substantially lower probability of mail return for 2010. Replacement forms were associated with a significant increase in mail returns, while bilingual forms were not (though the PAHUE DMAs may not have included enough Spanish-speaking households to provide a good sample to identify their effects). Several household demographic characteristics had significant effects in the expected direction. For example, married couple households had relatively high rates of return, renters had low rates relative to owners, and the age of the householder was consistently positively associated with the likelihood of a mail return. The partnership activity scale had no significant relationship to MRRs, though this may be due to the difficulty of quantifying geographic variation in the effectiveness of partnership activity rather than a lack of effect.

Thus, while we find that there are many factors that significantly affect mail return rates, the estimated PAHUE effect is still not significantly different from zero. For the overall sample, differences in the characteristics we can measure cannot explain the lack of evidence for a positive treatment effect. Comparing households that are similar along the dimensions included in the model, there is no statistical evidence that the households receiving the heavy-up treatment had higher mail return rates.

4.6.3 Logistic Model Results by Race/Ethnicity and Audience Segment

We now turn to examining different racial/ethnic groups and different audience segments, to see if we find any evidence of significant effects for some sub-populations. In doing so we use essentially the same model for each group.[13] In general, the odds ratios estimated for the demographic, housing, and survey variables do not vary substantially across groups, so we present only the estimated treatment effects for these specifications in the tables. See Appendix D for the complete results.

Table 35 presents estimated effects for three groups that are defined by the race and ethnicity of the householder: non-Hispanic Whites, non-Hispanic Blacks, and Hispanics. Again, we find no significant difference in MRRs associated with higher GRPs.

Table 35. Logistic Model of Mail Return Rate by Race/Ethnicity of Householder

Group	Odds Ratio for Log(GRP)	95% Confidence Interval
Non-Hispanic Whites	1.003	[0.971,1.036]
Non-Hispanic Blacks	0.956	[0.900,1.015]
Hispanics	0.903	[0.815,1.001]

Notes: The estimates presented here come from models with the same set of controls listed in the previous table. See Appendix Table D1 for the complete set of results.
Source: 2010 Mail Response/Return Assessment data for PAHUE areas.

[13] In the models presented here, our measure of exposure adds together GRPs from the DM plan and any targeted audience plan for the householder's race or ethnic group. Using this combined measure in our models implicitly assumes that the DM and targeted audience plans have similar effects on return rates. But it is possible, for example, that the BAP is more salient than the DM plan for Black householders and as a result has a larger effect on their MRRs. We explored models that included separate measures of GRPs from the DM plan and BAP, thereby allowing effects of the two plans to differ. (The Hispanic Audience Plan did not include local media buys, so it is not possible to explore such differences.) However, we decided to use the combined measure in this report based on several factors. First, the levels of BAP GRPs were not part of the PAHUE experiment and therefore were not manipulated experimentally across test and control DMAs. Consequently, any BAP effect detected could reflect correlation of GRPs with unmeasured factors that were outside of our control. Consistent with this, in the models explored we found significant effects from the BAP on the non-Hispanic/non-Black audience, potentially indicating a correlation between the BAP and unmeasured factors in the DMAs. Second, with only four pairs having local media buys in the BAP, we effectively have very few data points with which to distinguish differences in the effects of the two plans. Finally, although not under the control of our test, several pairs had large test/control differences in the BAP GRP delivery and some of the differences ran counter to the test/control DM heavy-up (e.g. the *control* site of Columbus received 6151 BAP GRPS while the *test* site of Savannah received only 4154 BAP GRPs).

In Table 36, we apply the logistic model to sub-populations defined by the primary audience segment for each tract. Here we do find a significant effect for one sub-population: Mobile singles. This group has roughly a 66 percent mail return rate on average. The point estimate from the model implies that, starting from 66 percent as a base, a 50 percent increase in GRPs would increase the mail return rate by roughly one percentage point.

Table 36. Logistic Model of Mail Return Rate by Audience Segment

Segment	Odds Ratio for Log(GRP)	95% Confidence Interval
All around average owner/renter	0.997	[0.944,1.053]
Economically disadvantaged owner/renter	0.971	[0.877,1.076]
Mobile singles	1.240*****	[1.025,1.500]
Advantaged home-owners	1.082	[0.984,1.189]

Notes: Audience segment is determined by the characteristics of a household's tract of residence, as measured in 2000. The estimates presented here come from models with the same set of controls listed in Table 34. See Appendix Table D2 For the complete set of results.
p-values: *p<0.05, ** p<0.01, *** p<0.001
Source: 2010 Mail Response/Return Assessment data for PAHUE areas.

4.6.4. Nonresponse Follow-Up Proxies

Households that failed to complete and return a census form by mail were followed-up by Census enumerators who attempted to obtain the information during a personal visit. If an enumerator could not make contact or gain cooperation, they attempted to obtain the information from a knowledgeable proxy such as a neighbor. A final hypothesis around the heavy-up experiment was whether the extra advertising might affect cooperation with enumerators. While the experiment did not alter the dosage of advertising during the NRFU phase, there was still reason to believe the extra advertising during the awareness and motivation phases might have facilitated a higher level of cooperation in areas that received more ads.

However, we found no evidence for this hypothesis. Overall, the difference in NRFU proxy response between treatment and control sites was less than one percent (5.9 percent in the control sites versus 6.6 percent in the treatment – see Appendix E). In some pairs the difference was a bit larger and in one pair it was in the other direction. Overall, however, we conclude the heavy-up had no effect (positive or negative) regarding cooperation with enumerators.

4.7 Timing of Mailback

One potential impact from the paid media is on the timing of mailback of census forms. Advertising frequently is used as a reminder to the members of the target audience to take a desired action – in this case, return their census form. The reminder helps to break the inertia imposed by the time demands of everyday life. The frequency of delivering the reminder is potentially connected to the length of the interval before action is finally taken at the individual

level. In this scenario, we would expect the treatment DMAs to respond more quickly than the control DMAs. Among other benefits, receiving a form sooner reduces the risk of cases ending up in the NRFU pool and helps the agency more accurately plan for the size and scope of the NRFU operation.

One way to assess the impact of a communications program on the length of time to obtain an action is to look at the observed cumulative distributions of populations being compared. If a group is taking action more quickly, then its cumulative distribution, at a given point in time, will be higher than the comparison population. That is, at a given point in time, the population that is acting more quickly will have returned a greater share of their forms. Similarly, the group that is acting more quickly will reach a given response level at an earlier date than the comparative population. In the context of the PAHUE, we would expect the cumulative distribution for the treatment DMAs to be above and to the left of that for the control DMAs. To measure timing, we count the cumulative number of forms received by a given date and divide by the total number received by April 19, i.e. total returns before the NRFU Cut-off. After April 19, the timing of returns would be less likely to have been affected by differences in paid media.

Analyzing the timing of mail returns from the 2010 Census for the PAHUE DMAs is complicated by logistical factors. In particular, the timing of mail returns was obtained from their check-in dates at several processing centers. Forms from the different members of a given pair of DMAs in the test were often processed by different centers. The members may be at substantially different distances to the processing centers which may affect the time it takes for returned forms to arrive. Table 37 gives the PAHUE DMA pairs and their associated processing centers.

Table 37. Processing Centers for PAHUE DMAs

Treatment DMA	Processing Center	Control DMA	Processing Center
Savannah, GA	Baltimore	Columbia, SC	Baltimore
Flint/Saginaw, MI	Jeffersonville	Toledo, OH	Jeffersonville
Montgomery, AL	Baltimore	Jackson, MS	Phoenix
Baton Rouge, LA	Phoenix	Augusta, GA	Baltimore
Shreveport, LA	Phoenix	Tallahassee, FL	Baltimore
Odessa, TX	Phoenix	Lubbock, TX	Phoenix
Erie, PA	Baltimore	Joplin, MO	Jeffersonville and Phoenix
Jacksonville, FL	Baltimore	Little Rock, AR	Phoenix

We attempted to work around this logistical issue by aligning the “start date” when census forms began to arrive in “reasonable” volume at the processing center. Formally, Day 1 was defined to be the first day with at least 2 percent of the eventual returns checked in. The range of actual dates for Day 1 was from March 18 to March 22, 2010. The range of the share of mail returns checked in on Day 1 is 3.9 percent to 18.6 percent. The graphs of the mail returns checked in by day since the initial check-ins are included in Appendix F.

An inspection of the graphs shows the possibility that the returns may have arrived slightly more quickly from treatment DMAs. Three pairs (Savannah/Columbia, Flint-Saginaw/Toledo, and Jacksonville/Little Rock) show a profile with the treatment slightly above the control. Three

pairs (Montgomery/Jackson, Shreveport/Tallahassee, and Odessa/Lubbock) show essentially no difference. One pair, Erie/Joplin, shows difference in favor of the treatment for a portion of the time period. One pair shows a strong reversal with the returns from the control site, Augusta, arriving substantially more quickly than from the treatment site, Baton Rouge.

Given outcomes of the eight pairs, it is unlikely that a formal statistical analysis would indicate a significant reduction in response times between the treatment and control DMAs. We did not pursue modeling the shapes of the cumulative return curves in order to do formal comparison of the timing of form return.

In the three DMAs where the cumulative distribution of responses from the treatment DMA is shifted to the left of cumulative distribution from the control DMA, the size of the shift is between 1 day and 3 days. An analysis of the operational benefit from this level of "pull forward" of returns could decide whether to pursue a deeper probe into factors that may have affected the timing of returns across the PAHUE pairs.

4.8 Cost/Benefit of the Heavy-Up

One rationale for exploring a heavy-up media schedule was to examine whether the extra investment in media was cost-justified through increased MRR and reduced NRFU. The Census Bureau has estimated that each NRFU case costs $33.65 (see Jackson, et al., forthcoming). The mail-back universe in PAHUE treatment DMAs is approximately 2.55 million households. The DraftFCB post-buy analysis puts the gross level of incremental media buy for the PAHUE at approximately $1.30 million. Comparing the actual level of PAHUE spend to the estimate NRFU savings in PAHUE sites implies that the PAHUE would break-even if it increased MRR by about 38,600 households or 1.5 percentage points.

Calculating the incremental media investment required to implement the PAHUE more broadly requires reference to the national DM media plan. This plan made substantial use of media properties, like network programming, that provide economies of scale in reaching the audience. Taking the DM national plan and pro-rating the cost to the 2.37 percent of US households in the PAHUE treatment DMAs implies an incremental spend of $1.42 million. Using the same analysis approach in the paragraphs above but with a spend of $1.42 million, the PAHUE would break-even if it could increase MRR by about 1.7 percentage points on a broader scale.

As noted in section 4.6, the overall observed MRR is slightly higher in the control DMAs than in the treatment DMAs. When analyzing sub-populations, in most cases, the statistical comparison concludes that any difference in MRR is statistically indistinguishable from zero. Therefore, in this test, the cost of incremental media in treatment DMAs over and above the base plan was not justified through an increase in mail returns. This is not an assertion in any way that the base plan was not effective. As noted above, the overall effect of the ICC was to markedly increase awareness and knowledge of the Census. The media levels deployed in this test raise the possibility that both treatment and control sites were at saturation levels.

5. Recommendations

As we frame the recommendations from this test, we should recall that this is the first test of its kind implemented with a Census Bureau paid media campaign. Census Advisory Committees,

academic experts, and program evaluation experts have recommended that this type of controlled in-market test be undertaken to better understand the impact from the integrated communications campaign. One implicit objective for this test was to probe what worked and what did not when implementing a test in the open market during a Census. Consequently, several recommendations relate to how to improve the design and implementation of communication tests in the future.

Secondly, this test is not projectable to the nation or other subsets of DMAs. The test was not designed to make broad inferences about what worked and what did not in the ICC. The test did identify some areas and hypotheses for further exploration. These will be pointed out.

- Continue in-market testing of variables subject to Census Bureau control in the communication strategy, but design the tests differently.
    - Given the evidence in this test of saturation, design media weight tests that investigate reduced media levels. Do not repeat a heavy-up only design.
        - Reduced media schedules could be tested in the open market, but we understand the difficulties with approving such a test.
        - Investigate use of technical solutions to reduce media exposure in a very controlled way to a limited number of known households. When the households receiving reduced media level are known in advance, they could receive special monitoring and NRFU to ensure that they are fully counted. Determine if Behaviorscan[14] or a similar testing service can provide a practical approach to testing lowered paid media weight for the 2020 Census or, better yet, for tests ahead of 2020.
    - Design tests to examine the relative impact from the large-scale allocations to communication strategies in the Integrated Campaign. In particular, design factorial experiments that vary the levels of partnership and earned media simultaneously with each other and with Paid Media.
    - Design tests to include other variables in the paid media mix:
        - The mix of media vehicles, such as TV, radio, print, online, social media, new/emerging media – especially the relative mix of new versus traditional media.
        - The mix of targeted campaigns versus more general, Diverse Mass, campaigns
        - The start date, modulation of GRP levels, and flighting (use of hiatus periods) for the campaign.
    - Design tests that use moderately sized DMAs not receiving spending through local targeted advertising plans. It will be easier to control the media delivery in these DMAs to the design specification. When using this type of DMA, the tests will naturally focus more on households already inclined to participate and for which the paid and earned media campaigns are expected to "close the deal." At the same time, these DMAs would still allow for exploration of the interactions between paid media and partnership programs for target populations.

[14] Behaviorscan is a testing service that can selectively add or delete advertising to sampled households in several smaller DMAs.

- The large differences in mail return rates between matched DMAs pairs and from prior experience point to the presence of unmeasured, uncontrolled covariates in this test.
    - Design measurements to more precisely understand the level and impact of partnerships in test locations. These measurements would obtain information on the geographic scope, the number of people reached, and the timing/frequency of the program. It is expected that this measurement system would be parallel to the data gathered for managing the partnership program.
    - Design measurements that would more precisely understand the level and valence of earned media in test locations.
    - To better understand the number of pairs needed for testing, analyze variation in mail return rates among DMAs in the clusters of similar DMAs from which the "matched markets" were drawn. The analysis should include the decade-to-decade differences in addition to DMA-to-DMA differences.

- Any future in-market testing should include a tracking survey as part of the design. The CICPE survey implemented in the PAHUE markets provided extremely valuable data on the change in key measures like awareness and knowledge of the Census over time.

- Use measures for advertising exposure and frequency of exposure which do not depend on self reports.
    - Continue to obtain GRP and impressions data for the media used at the local market level. Plan in advance to obtain this information so that the contracting agencies can build appropriate media data collection into their resource plans.
    - The self-reported measures of advertising exposure and frequency in this test do not align with the large differences in measured exposure between treatment and control DMAs. Explore alternative methods with service providers and individuals who specialize in analyzing advertising effects and tests.
    - Determine whether surveys can be an effective tool to model the impact on mail response of an ICC. That is, determine whether the role of surveys should primarily be to understand the campaign's impact on message reception, awareness, knowledge, attitudes, and intention to act, but stop short of testing the goals for behavioral change.

- Given evidence of paid media saturation in the PAHUE, consider reducing the national DM campaign to free up resources for local partnership activity, target audience plans, and non-advertising uses.

- Given the evidence of an enhanced positive communication impact among the Black audience from "Miss Maybelle," continue to use target audience plans. Consider how these ads and media timing can be used to rapidly increase awareness, knowledge, and intention to respond among target audiences. Plan exploratory research, hopefully ahead of the 2020 Census, to understand the role of paid media relative to earned media and partnership in driving MRR and enumeration response from the target audiences.

6. Related Assessments, Evaluations, or Experiments

The 2010 CICPE conducted by NORC contains a comprehensive evaluation of the entire 2010 Census ICC as it relates to changes in attitudes, opinions, awareness and knowledge over time. It also examines the relationship between campaign exposure and mail return behavior. This report contains a detailed methodology of the survey data collection used in the PAHUE and also a description of the Census operational data extracts used to analyze the survey participant's mail return rates.

The 2010 Census ICC Assessment (Paid Media; Census in Schools; Partnerships; and Earned Media) reports on the successes and shortcomings of the different components of the campaign from an operational standpoint. The report "2010 Census Program for Evaluations and Experiments: The Mail Response/Return Rate Assessment" contains the overall results for the 2010 Census mail response and mail return rates. The authors of this report produced the data extract used to conduct the PAHUE mail return rate modeling.

7. Acknowledgements

We acknowledge and thank Earl Letourneau for producing the PAHUE 2010 Census response data extract necessary to compute the mail return models. We also thank Dennis Stoudt and his staff for providing the data necessary to analyze the NRFU proxy data. We thank Burton Reist and Frank Vitrano for help promoting and advocating the PAHUE. We also thank Laura Sewell and Alexandra Figueroa for managing the heavy-up media and post-buy audit contracts. Finally, we thank our critical reviewers: Mike Bentley, Lucia Foster, Tasha Boone, Adrienne Oneto, Fred Meier, and Jason Machowski for their review and comments on this report.

8. References

Academic Assessment Panel (2009). "Academic Assessment Panel Final Report". A report submitted to the U.S. Census Bureau, September 8, 2009. http://2010.census.gov/news/releases/operations/independent-panel-commends-the-2010-census-paid-media-plan.html.

American Association for Public Opinion Research (2011). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 7th edition. AAPOR.

Bates, N. and Mulry, M. (2011). "Using a Geographic Segmentation to Understand, Predict, and Plan for Census and Survey Mail Nonresponse." *Journal of Official Statistics*, 27:601-618.

Bishop, G.F., Oldendick, R.W., Tuchfarber, A.J., and Bennett, S.E. 1980. "Pseudo-Opinions on Public Affairs." *Public Opinion Quarterly* 44:198-209

Datta, R., Yan, T., Evans, D., Pedlow, S., Spencer, B., and Bautista, R. (2011). "Final Report: 2010 Census Integrated Communications Program Evaluation (CICPE). National Opinion Research Center at the University of Chicago. March 13, 2012.

Datta, R., Hepburn, P., Yan., T. and Evans, D. (2011). "Have you Really Seen this Ad?: Using Confirmed Awareness to Correct Estimates of Self-Reported Exposure." A paper presented at the Joint Statistical Meetings of the American Statistical Association, Miami, FL, July 30-August 4.

Edward, S. and Wilson, M. (2004). Census 2000 Testing, Experimentation and Evaluation Program Topic Report No. 6, TR-6, Evaluation of the Census 2000 Partnership and Marketing Program.

Jackson, G., Heimel, S. Walker S. and Winder, S. (2011). "2010 Census Nonresponse Followup Operations (NRO) Assessment." 2010 Census Program of Evaluations and Experiments, U.S. Census Bureau, *forthcoming report.*

Letourneau, E. (2011). "2010 Census Program for Evaluations and Experiments: The Mail Response/Return Rate Assessment", *forthcoming report.*

Lodish, L., Abraham, M., Kalmenson, S., Livelsberger, J., Lebetkin, B., Richardson, B., Stevens, M. (1995). "How T.V. Advertising Works: A Meta-Analysis of 389 Real World Split Cable T.V. Advertising Experiments." Journal of Marketing Research, Vol XXXII (May 1995), 125-139.

Miller, P. and Walejko, G. (2010). "Tracking Study of Attitudes and Intention to Participate in the 2010 Census." Technical Report Contract YA1323-09-CQ-0032. December 10, 2010.

National Research Council (2004). The 2000 Census: Counting Under Adversity. Panel to Review the 2000 Census. Constance F. Citro, Daniel L. Cork and Janet L. Norwood, eds. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, D.C.: The National Academies Press.

U.S. Census Bureau (2009). "Recommendations made by the Census Advisory Committee of Professional Associations (CACPA) at the meetings held April 16-17, 2009." Internal U.S. Census Bureau Memorandum from Jeri Green, Chief Advisory Committee Office, June 19, 2009.

Wolter, K., Calder, B., Malthouse, E. Murphy, S. Pedlow, S., and Porras, J. (2002). Partnership and Marketing Program Evaluation. Final Report. Census 2000 Evaluation, July 17, 2002.

9. Appendices

Appendix A: DMA Matched-Pair Selection Criteria Examples

Table A. DMA Matched-Pair Selection Criteria Examples

DMA Name	Number of DMA Households	Mean Hard-to-Count Score	Census 2000 Mail Return Rate	Census 2000 Percent Hispanic Population	Census 2000 Percent non-Hispanic White Population	Census 2000 Percent Black population
Columbia, SC	378,823	35	70.0%	2.1%	57.7%	38.1%
Savannah	300,094	43	68.6%	4.0%	60.6%	33.2%
Toledo	443,470	23	78.4%	4.7%	85.1%	8.1%
Flint-Saginaw-Bay City	473,197	26	80.5%	3.2%	82.3%	11.8%
Jackson, MS	354,187	43	74.5%	1.3%	51.3%	46.4%
Montgomery-Selma	268,406	40	68.3%	1.0%	54.8%	42.9%
Augusta-Aiken	265,185	42	70.4%	2.2%	57.5%	38.0%
Baton Rouge	299,417	38	69.1%	1.6%	62.3%	34.1%
Tallahassee-Thomasville	260,553	40	71.4%	3.6%	62.3%	31.9%
Shreveport	424,123	42	72.8%	3.8%	63.6%	30.2%
Lubbock	163,482	43	72.4%	34.1%	57.9%	6.4%
Odessa-Midland	145,314	45	73.0%	39.4%	54.0%	4.9%
Joplin-Pittsburg	171,494	30	78.0%	2.9%	90.2%	1.3%
Erie	164,142	25	82.5%	1.7%	92.1%	4.5%
Little Rock-Pine Bluff	579,035	36	74.9%	2.1%	76.9%	18.8%
Jacksonville	615,459	34	71.6%	3.7%	71.7%	21.1%

Source: 2010 Mail Response/Return Assessment data for PAHUE areas.

Appendix B: Local and National Media Buys by Type of Media

Table B1. Diverse Mass Audience Actualized Local Media Buys by DMA

	Net Spending					
DMA Name	**Print**	**Radio**	**Television**	**Digital**	**Out-of-Home**	**Total**
Columbia, SC	$2,261	$0	$0	$0	$2,469	**$4,730**
Savannah	$4,739	$20,651	$101,645	$17,084	$2,469	**$146,588**
Toledo	$0	$765	$0	$0	$2,272	**$3,037**
Flint-Saginaw-Bay City	$16,072	$16,423	$110,111	$17,084	$1,136	**$160,827**
Jackson, MS	$17,348	$1,573	$0	$0	$1,333	**$20,254**
Montgomery-Selma	$63,434	$22,895	$65,538	$17,084	$2,583	**$171,534**
Augusta-Aiken	$0	$850	$0	$0	$1,333	**$2,183**
Baton Rouge	$37,259	$30,745	$146,208	$17,084	$2,469	**$233,766**
Tallahassee-Thomasville	$50,701	$2,479	$1,560	$0	$2,583	**$57,322**
Shreveport	$88,572	$20,103	$130,014	$17,084	$1,333	**$257,105**
Lubbock	$3,400	$1,341	$3,188	$0	$2,560	**$10,490**
Odessa-Midland	$6,862	$12,735	$66,895	$17,084	$1,333	**$104,909**
Joplin-Pittsburg	$266	$0	$0	$0	$1,333	**$1,600**
Erie	$5,945	$10,468	$75,974	$17,084	$1,136	**$110,607**
Little Rock-Pine Bluff	$116,006	$1,318	$0	$0	$2,697	**$120,020**
Jacksonville	$88,243	$70,248	$176,239	$17,084	$2,469	**$354,283**
Total	**$501,108**	**$212,592**	**$877,372**	**$136,672**	**$31,512**	**$1,759,255**

Source: 2010 Mail Response/Return Assessment data for PAHUE areas.

Table B2. Diverse Mass Audience Actualized National Media Buys by DMA

	Net Spending					
DMA Name	**Print**	**Radio**	**Television**	**Cinema**	**Digital**	**TOTAL**
Columbia, SC	$14,062	$10,242	$151,671	$3,012	$29,587	**$208,574**
Savannah	$11,347	$8,264	$122,386	$2,431	$23,874	**$168,302**
Toledo	$14,913	$10,861	$160,850	$3,195	$31,377	**$221,197**
Flint-Saginaw-Bay City	$16,170	$11,776	$174,400	$3,464	$34,021	**$239,830**
Jackson, MS	$11,874	$8,648	$128,068	$2,544	$24,983	**$176,116**
Montgomery-Selma	$8,632	$6,287	$93,101	$1,849	$18,161	**$128,030**
Augusta-Aiken	$9,037	$6,582	$97,471	$1,936	$19,014	**$134,040**
Baton Rouge	$11,550	$8,412	$124,571	$2,474	$24,300	**$171,307**
Tallahassee-Thomasville	$9,888	$7,202	$106,650	$2,118	$20,805	**$146,663**
Shreveport	$13,617	$9,917	$146,863	$2,917	$28,649	**$201,962**
Lubbock	$5,593	$4,073	$60,319	$1,198	$11,767	**$82,949**
Odessa-Midland	$5,066	$3,689	$54,636	$1,085	$10,658	**$75,135**
Joplin-Pittsburg	$5,511	$4,014	$59,444	$1,181	$11,596	**$81,747**
Erie	$5,511	$4,014	$59,444	$1,181	$11,596	**$81,747**
Little Rock-Pine Bluff	$19,898	$14,492	$214,612	$4,263	$41,865	**$295,129**
Jacksonville	$23,951	$17,443	$258,321	$5,131	$50,391	**$355,237**
Total	**$186,621**	**$135,916**	**$2,012,808**	**$39,978**	**$392,643**	**$2,767,965**

Source: 2010 Mail Response/Return Assessment data for PAHUE areas.

Table B3. Combined Diverse Mass, Black, and Hispanic Audience Plan Local Media Buys

	Net Spending					
DMA Name	**Print**	**Radio**	**Television**	**Digital**	**Out-of-Home**	**Total**
Columbia, SC	$20,419	$28,375	$0	$0	$20,559	**$69,353**
Savannah	$24,123	$43,158	$101,645	$17,084	$27,669	**$213,679**
Toledo	$16,028	$36,975	$11,377	$0	$2,272	**$66,651**
Flint-Saginaw-Bay City	$16,072	$23,967	$147,507	$17,084	$1,136	**$205,767**
Jackson, MS	$60,093	$44,783	$0	$0	$6,013	**$110,889**
Montgomery-Selma	$63,934	$73,812	$65,538	$17,084	$8,523	**$228,891**
Augusta-Aiken	$8,183	$41,645	$0	$0	$5,233	**$55,061**
Baton Rouge	$51,775	$58,592	$146,208	$23,087	$27,258	**$306,920**
Tallahassee-Thomasville	$79,966	$22,233	$39,239	$0	$15,183	**$156,620**
Shreveport	$90,862	$61,906	$145,241	$17,084	$5,533	**$320,625**
Lubbock	$31,222	$14,705	$3,188	$0	$2,560	**$51,675**
Odessa-Midland	$13,832	$19,159	$78,093	$17,084	$1,333	**$129,500**
Joplin-Pittsburg	$266	$0	$0	$0	$1,333	**$1,600**
Erie	$5,945	$10,468	$75,974	$17,084	$1,136	**$110,607**
Little Rock-Pine Bluff	$127,651	$5,958	$0	$5,641	$2,697	**$141,946**
Jacksonville	$116,878	$86,266	$176,239	$17,084	$2,469	**$398,936**
Total	**$727,247**	**$572,000**	**$990,248**	**$148,315**	**$130,911**	**$2,568,721**

Source: 2010 Mail Response/Return Assessment data for PAHUE areas.

Table B4. Combined Diverse Mass, Black, and Hispanic Audience Plan Actualized National Media Buys

	Net Spending					
DMA Name	**Print**	**Radio**	**Television**	**Cinema**	**Digital**	**TOTAL**
Columbia, SC	$22,575	$33,720	$226,095	$3,012	$35,629	**$321,031**
Savannah	$17,034	$23,602	$172,901	$2,431	$28,324	**$244,292**
Toledo	$17,848	$18,239	$188,152	$3,195	$34,315	**$261,749**
Flint-Saginaw-Bay City	$19,302	$20,024	$202,686	$3,464	$36,711	**$282,187**
Jackson, MS	$20,343	$32,325	$201,372	$2,544	$30,611	**$287,194**
Montgomery-Selma	$14,458	$22,569	$143,537	$1,849	$22,039	**$204,452**
Augusta-Aiken	$14,379	$21,365	$144,060	$1,936	$22,746	**$204,486**
Baton Rouge	$18,064	$26,384	$181,509	$2,474	$34,919	**$263,351**
Tallahassee-Thomasville	$15,047	$21,215	$152,249	$2,118	$24,723	**$215,353**
Shreveport	$20,646	$28,952	$209,123	$2,917	$34,057	**$295,694**
Lubbock	$8,955	$10,953	$95,245	$1,198	$17,017	**$133,367**
Odessa-Midland	$8,836	$11,243	$94,163	$1,085	$16,735	**$132,064**
Joplin-Pittsburg	$5,930	$4,931	$63,656	$1,181	$12,178	**$87,877**
Erie	$5,991	$5,247	$63,847	$1,181	$12,044	**$88,310**
Little Rock-Pine Bluff	$26,439	$32,183	$272,596	$4,263	$52,563	**$388,044**
Jacksonville	$33,428	$42,479	$343,723	$5,131	$58,436	**$483,198**
Total	**$269,276**	**$355,430**	**$2,754,915**	**$39,978**	**$473,050**	**$3,892,649**

Source: 2010 Mail Response/Return Assessment data for PAHUE areas.

Appendix C: DMA Mail Return Rates by Race/Ethnicity and Audience Segment

Table C1. Mail Return Rates by DMA by Race/Ethnicity of Householder

Treatment	DMA	Race/Ethnicity	Mail Return Rate
Control	COLUMBIA, SC	Non-Hisp Whites	81.90%
Treatment	SAVANNAH	Non-Hisp Whites	77.60%
Control	COLUMBIA, SC	Non-Hisp Blacks	73.80%
Treatment	SAVANNAH	Non-Hisp Blacks	69.40%
Control	COLUMBIA, SC	Hispanics	58.20%
Treatment	SAVANNAH	Hispanics	55.40%
Control	TOLEDO	Non-Hisp Whites	80.90%
Treatment	FLINT-SAGINAW-BAY CITY	Non-Hisp Whites	85.10%
Control	TOLEDO	Non-Hisp Blacks	63.10%
Treatment	FLINT-SAGINAW-BAY CITY	Non-Hisp Blacks	72.80%
Control	TOLEDO	Hispanics	62.70%
Treatment	FLINT-SAGINAW-BAY CITY	Hispanics	72.00%
Control	JACKSON, MS	Non-Hisp Whites	77.60%
Treatment	MONTGOMERY (SELMA)	Non-Hisp Whites	78.60%
Control	JACKSON, MS	Non-Hisp Blacks	66.90%
Treatment	MONTGOMERY (SELMA)	Non-Hisp Blacks	66.50%
Control	JACKSON, MS	Hispanics	50.70%
Treatment	MONTGOMERY (SELMA)	Hispanics	49.90%
Control	AUGUSTA	Non-Hisp Whites	79.30%
Treatment	BATON ROUGE	Non-Hisp Whites	73.90%
Control	AUGUSTA	Non-Hisp Blacks	71.40%
Treatment	BATON ROUGE	Non-Hisp Blacks	61.00%
Control	AUGUSTA	Hispanics	59.50%
Treatment	BATON ROUGE	Hispanics	55.70%
Control	TALLAHASSEE-THOMASVILLE	Non-Hisp Whites	76.10%
Treatment	SHREVEPORT	Non-Hisp Whites	76.10%
Control	TALLAHASSEE-THOMASVILLE	Non-Hisp Blacks	66.20%
Treatment	SHREVEPORT	Non-Hisp Blacks	65.80%
Control	TALLAHASSEE-THOMASVILLE	Hispanics	58.60%
Treatment	SHREVEPORT	Hispanics	57.60%
Control	LUBBOCK	Non-Hisp Whites	79.80%
Treatment	ODESSA-MIDLAND	Non-Hisp Whites	77.30%
Control	LUBBOCK	Non-Hisp Blacks	66.00%
Treatment	ODESSA-MIDLAND	Non-Hisp Blacks	62.30%
Control	LUBBOCK	Hispanics	66.20%
Treatment	ODESSA-MIDLAND	Hispanics	65.40%
Control	JOPLIN-PITTSBURG	Non-Hisp Whites	77.20%
Treatment	ERIE	Non-Hisp Whites	83.30%

Control	JOPLIN-PITTSBURG	Non-Hisp Blacks	59.50%
Treatment	ERIE	Non-Hisp Blacks	60.30%
Control	JOPLIN-PITTSBURG	Hispanics	55.20%
Treatment	ERIE	Hispanics	67.20%
Control	LITTLE ROCK-PINE BLUFF	Non-Hisp Whites	76.70%
Treatment	JACKSONVILLE	Non-Hisp Whites	75.60%
Control	LITTLE ROCK-PINE BLUFF	Non-Hisp Blacks	61.30%
Treatment	JACKSONVILLE	Non-Hisp Blacks	64.40%
Control	LITTLE ROCK-PINE BLUFF	Hispanics	50.10%
Treatment	JACKSONVILLE	Hispanics	63.80%

Source: 2010 Mail Response/Return Assessment data for PAHUE areas.

Table C2. Mail Return Rates by DMA by 2010 Communication Campaign Audience Segment

Treatment	DMA	Audience Segment	Mail Return Rate
Control	COLUMBIA, SC	All around avg I (homeowner skew)	78.40%
Treatment	SAVANNAH	All around avg I (homeowner skew)	75.00%
Control	COLUMBIA, SC	All around avg II (renter skew)	75.10%
Treatment	SAVANNAH	All around avg II (renter skew)	71.90%
Control	COLUMBIA, SC	Econ disadvantaged I (homeowner skew)	75.50%
Treatment	SAVANNAH	Econ disadvantaged I (homeowner skew)	71.20%
Control	COLUMBIA, SC	Econ disadvantaged II (renter skew)	69.30%
Treatment	SAVANNAH	Econ disadvantaged II (renter skew)	57.90%
Treatment	SAVANNAH	Ethnic enclave I (homeowner skew)	61.00%
Control	COLUMBIA, SC	Mobile singles	68.90%
Treatment	SAVANNAH	Mobile singles	66.30%
Control	COLUMBIA, SC	Advantaged homeowners	82.80%
Treatment	SAVANNAH	Advantaged homeowners	81.60%
Control	TOLEDO	All around avg I (homeowner skew)	79.10%
Treatment	FLINT-SAGINAW-BAY CITY	All around avg I (homeowner skew)	83.30%
Control	TOLEDO	All around avg II (renter skew)	74.40%
Treatment	FLINT-SAGINAW-BAY CITY	All around avg II (renter skew)	80.40%
Control	TOLEDO	Econ disadvantaged I (homeowner skew)	68.80%
Treatment	FLINT-SAGINAW-BAY CITY	Econ disadvantaged I (homeowner skew)	75.30%
Control	TOLEDO	Econ disadvantaged II (renter skew)	64.80%
Treatment	FLINT-SAGINAW-BAY CITY	Econ disadvantaged II (renter skew)	72.20%
Control	TOLEDO	Mobile singles	65.80%
Treatment	FLINT-SAGINAW-BAY CITY	Mobile singles	74.40%
Control	TOLEDO	Advantaged homeowners	83.40%
Treatment	FLINT-SAGINAW-BAY CITY	Advantaged homeowners	86.70%
Control	JACKSON, MS	All around avg I (homeowner skew)	73.50%
Treatment	MONTGOMERY (SELMA)	All around avg I (homeowner skew)	74.60%
Control	JACKSON, MS	All around avg II (renter skew)	68.30%
Treatment	MONTGOMERY (SELMA)	All around avg II (renter skew)	68.00%
Control	JACKSON, MS	Econ disadvantaged I (homeowner skew)	68.20%
Treatment	MONTGOMERY (SELMA)	Econ disadvantaged I (homeowner skew)	67.30%
Control	JACKSON, MS	Econ disadvantaged II (renter skew)	65.40%
Treatment	MONTGOMERY (SELMA)	Econ disadvantaged II (renter skew)	70.10%
Control	JACKSON, MS	Mobile singles	60.00%
Treatment	MONTGOMERY (SELMA)	Mobile singles	65.10%
Control	JACKSON, MS	Advantaged homeowners	78.00%
Treatment	MONTGOMERY (SELMA)	Advantaged homeowners	79.10%
Control	AUGUSTA	All around avg I (homeowner skew)	77.50%
Treatment	BATON ROUGE	All around avg I (homeowner skew)	69.70%
Control	AUGUSTA	All around avg II (renter skew)	69.70%
Treatment	BATON ROUGE	All around avg II (renter skew)	67.20%

Control	AUGUSTA	Econ disadvantaged I (homeowner skew)	73.00%
Treatment	BATON ROUGE	Econ disadvantaged I (homeowner skew)	64.80%
Control	AUGUSTA	Econ disadvantaged II (renter skew)	74.40%
Treatment	BATON ROUGE	Econ disadvantaged II (renter skew)	56.30%
Control	AUGUSTA	Ethnic enclave I (homeowner skew)	78.00%
Treatment	BATON ROUGE	Ethnic enclave I (homeowner skew)	64.00%
Treatment	BATON ROUGE	Mobile singles	61.20%
Control	AUGUSTA	Advantaged homeowners	78.80%
Treatment	BATON ROUGE	Advantaged homeowners	76.20%
Control	TALLAHASSEE-THOMASVILLE	All around avg I (homeowner skew)	74.20%
Treatment	SHREVEPORT	All around avg I (homeowner skew)	73.10%
Control	TALLAHASSEE-THOMASVILLE	All around avg II (renter skew)	72.20%
Treatment	SHREVEPORT	All around avg II (renter skew)	69.50%
Control	TALLAHASSEE-THOMASVILLE	Econ disadvantaged I (homeowner skew)	70.90%
Treatment	SHREVEPORT	Econ disadvantaged I (homeowner skew)	69.30%
Control	TALLAHASSEE-THOMASVILLE	Econ disadvantaged II (renter skew)	57.00%
Treatment	SHREVEPORT	Econ disadvantaged II (renter skew)	63.90%
Treatment	SHREVEPORT	Ethnic enclave I (homeowner skew)	73.60%
Control	TALLAHASSEE-THOMASVILLE	Mobile singles	64.00%
Treatment	SHREVEPORT	Mobile singles	62.40%
Control	TALLAHASSEE-THOMASVILLE	Advantaged homeowners	80.10%
Treatment	SHREVEPORT	Advantaged homeowners	79.20%
Control	LUBBOCK	All around avg I (homeowner skew)	77.00%
Treatment	ODESSA-MIDLAND	All around avg I (homeowner skew)	72.50%
Control	LUBBOCK	All around avg II (renter skew)	72.30%
Treatment	ODESSA-MIDLAND	All around avg II (renter skew)	70.80%
Control	LUBBOCK	Econ disadvantaged I (homeowner skew)	70.00%
Treatment	ODESSA-MIDLAND	Econ disadvantaged I (homeowner skew)	69.40%
Control	LUBBOCK	Econ disadvantaged II (renter skew)	62.40%
Treatment	ODESSA-MIDLAND	Econ disadvantaged II (renter skew)	66.40%
Control	LUBBOCK	Ethnic enclave I (homeowner skew)	72.90%
Treatment	ODESSA-MIDLAND	Ethnic enclave I (homeowner skew)	70.60%
Control	LUBBOCK	Mobile singles	68.60%
Treatment	ODESSA-MIDLAND	Mobile singles	67.50%
Control	LUBBOCK	Advantaged homeowners	82.60%
Treatment	ODESSA-MIDLAND	Advantaged homeowners	76.10%
Control	JOPLIN-PITTSBURG	All around avg I (homeowner skew)	76.80%
Treatment	ERIE	All around avg I (homeowner skew)	83.50%
Control	JOPLIN-PITTSBURG	All around avg II (renter skew)	72.70%
Treatment	ERIE	All around avg II (renter skew)	77.20%
Control	JOPLIN-PITTSBURG	Econ disadvantaged I (homeowner skew)	72.90%
Treatment	ERIE	Econ disadvantaged I (homeowner skew)	71.60%
Treatment	ERIE	Econ disadvantaged II (renter skew)	71.20%
Control	JOPLIN-PITTSBURG	Ethnic enclave I (homeowner skew)	68.10%

Treatment	ERIE	Mobile singles	71.10%
Control	JOPLIN-PITTSBURG	Advantaged homeowners	78.10%
Treatment	ERIE	Advantaged homeowners	86.10%
Control	LITTLE ROCK-PINE BLUFF	All around avg I (homeowner skew)	73.70%
Treatment	JACKSONVILLE	All around avg I (homeowner skew)	73.90%
Control	LITTLE ROCK-PINE BLUFF	All around avg II (renter skew)	70.40%
Treatment	JACKSONVILLE	All around avg II (renter skew)	69.10%
Control	LITTLE ROCK-PINE BLUFF	Econ disadvantaged I (homeowner skew)	68.40%
Treatment	JACKSONVILLE	Econ disadvantaged I (homeowner skew)	69.00%
Control	LITTLE ROCK-PINE BLUFF	Econ disadvantaged II (renter skew)	63.60%
Treatment	JACKSONVILLE	Econ disadvantaged II (renter skew)	70.70%
Control	LITTLE ROCK-PINE BLUFF	Ethnic enclave I (homeowner skew)	65.40%
Control	LITTLE ROCK-PINE BLUFF	Mobile singles	65.90%
Treatment	JACKSONVILLE	Mobile singles	64.40%
Control	LITTLE ROCK-PINE BLUFF	Advantaged homeowners	79.40%
Treatment	JACKSONVILLE	Advantaged homeowners	77.40%

Source: 2010 Mail Response/Return Assessment data for PAHUE areas.

Appendix D: Logistic Models by Race/Ethnicity and Audience Segment

Table D1. Logistic Model of Mail Return Rate by Race/Ethnicity Group

Odds ratios/[95% confidence intervals] Dep var = Valid mail return by 4/19/10	**Non-Hispanic Whites**		**Non-Hispanic Blacks**		**Hispanics**	
Log (GRP)	1.003		0.956		0.903	
	[0.971,1.036]		[0.900,1.015]		[0.815,1.001]	
Type of enumeration area						
Update/leave enumeration area	0.72	***	0.64	***	0.727	***
	[0.660,0.786]		[0.530,0.773]		[0.638,0.828]	
Military enumeration area	1.024		1		1.356	***
	[0.850,1.234]		[0.760,1.315]		[1.187,1.549]	
Urban update/leave enumeration area	0.625	***	0.618	***	0.698	***
	[0.561,0.695]		[0.542,0.704]		[0.596,0.818]	
Outcome of 2000 collection at this address						
Address not in 2000 census	0.811	***	0.833	***	0.962	
	[0.786,0.837]		[0.804,0.862]		[0.902,1.025]	
Enumerator completed return in 2000	0.616	***	0.656	***	0.725	***
	[0.601,0.632]		[0.644,0.668]		[0.689,0.764]	
Other outcome in 2000	0.718	***	0.712	***	0.769	***
	[0.700,0.736]		[0.666,0.760]		[0.698,0.849]	
Type of housing unit						
Multi-unit building	0.839	***	0.868	***	0.919	**
	[0.813,0.865]		[0.845,0.892]		[0.867,0.973]	
Trailer/mobile home/boat/RV, etc	0.741	***	0.786	***	0.626	***
	[0.712,0.771]		[0.772,0.800]		[0.569,0.689]	
Married couple	1.807	***	1.355	***	1.478	***
	[1.757,1.857]		[1.307,1.405]		[1.433,1.524]	
Household rents or occupies	0.6	***	0.696	***	0.592	***
without paying rent	[0.585,0.616]		[0.676,0.717]		[0.570,0.616]	
DMA pair for PAHUE						
Savannah/Columbia SC	1.174		1.402	**	1.054	
	[0.939,1.468]		[1.173,1.676]		[0.719,1.545]	
Flint-Saginaw/Toledo	1.478	*	1.253		1.209	
	[1.075,2.032]		[0.828,1.895]		[0.699,2.092]	
Selma/Jackson MS	0.998		1.13	**	0.815	
	[0.868,1.148]		[1.051,1.214]		[0.561,1.184]	
Baton Rouge/Augusta	0.975		1.053		1.01	
	[0.866,1.097]		[0.732,1.516]		[0.697,1.463]	
Shreveport/Tallahassee	1.066		1.186	**	1.081	
	[0.963,1.180]		[1.069,1.315]		[0.679,1.721]	
Odessa-Midland/Lubbock	1.194		1.027		1.046	
	[0.829,1.719]		[0.755,1.395]		[0.700,1.563]	
Erie/Joplin-Pittsburg	1.317		0.935		1.099	
	[0.965,1.799]		[0.849,1.030]		[0.582,2.076]	
Household of unrelated people	0.767	***	0.714	***	0.57	***
	[0.736,0.800]		[0.665,0.768]		[0.496,0.655]	

Table D1. Logistic Model of Mail Return Rate by Race/Ethnicity Group (continued)

Odds ratios/[95% confidence intervals] Dep var = Valid mail return by 4/19/10	**Non-Hispanic Whites**		**Non-Hispanic Blacks**		**Hispanics**	
Partnership activity scale (#1)	1.001		0.998		1.002	
	[0.989,1.013]		[0.992,1.004]		[0.996,1.007]	
Age of householder						
Householder age <=25	0.822	***	0.586	***	0.777	***
	[0.801,0.844]		[0.562,0.610]		[0.717,0.843]	
Householder age 41-64	1.52	***	1.763	***	1.648	***
	[1.482,1.560]		[1.716,1.811]		[1.598,1.701]	
Householder age 65+	3.067	***	2.616	***	2.883	***
	[2.904,3.238]		[2.476,2.763]		[2.740,3.034]	
Number of children age<18 in household						
1 child	0.918	***	1.058	**	1.049	
	[0.898,0.939]		[1.030,1.088]		[0.993,1.108]	
2 children	0.978	**	1.048	***	1.13	***
	[0.963,0.993]		[1.033,1.063]		[1.078,1.184]	
3-4 children	0.953	***	0.963	**	1.095	*
	[0.937,0.970]		[0.943,0.984]		[1.029,1.167]	
5-8 children	0.827	**	0.8	***	0.9	*
	[0.740,0.924]		[0.755,0.848]		[0.828,0.977]	
9+ children	0.759		0.558	*	1.726	
	[0.437,1.317]		[0.354,0.882]		[0.379,7.858]	
Number of residents in household						
2 people	1.074	***	1.268	***	1.267	***
	[1.052,1.095]		[1.240,1.297]		[1.198,1.341]	
3 people	0.913	***	1.226	***	1.051	
	[0.889,0.938]		[1.190,1.263]		[0.975,1.132]	
4 people	0.811	***	1.177	***	0.939	
	[0.781,0.841]		[1.139,1.216]		[0.855,1.030]	
5 people	0.663	***	1.153	***	0.819	***
	[0.632,0.694]		[1.113,1.193]		[0.769,0.874]	
9-12 people	0.602	***	1.373	***	0.894	*
	[0.550,0.659]		[1.304,1.447]		[0.824,0.970]	
13+ people	0.222	***	0.52	*	0.274	***
	[0.148,0.333]		[0.290,0.934]		[0.158,0.476]	
In area blanketed by replacement forms	1.231	***	1.212	***	1.196	***
	[1.152,1.315]		[1.148,1.280]		[1.128,1.269]	
In area where non-respondents received replacement form	1.141	**	1.153	***	1.06	*
	[1.073,1.214]		[1.103,1.205]		[1.018,1.104]	
Address is not residential	0.034	***	0.042	***	0.035	**
	[0.027,0.043]		[0.017,0.106]		[0.005,0.247]	
Initial form was bilingual					1.161	***
					[1.088,1.240]	

Note: Each column contains odds ratio with 95% confidence interval below it. Omitted categories: In mail-out/mail-back type of enumeration area, address had respondent return in 2000, owner, single-family house, Jacksonville/Little Rock DMA pair, single person household, householder aged 26-40.

p-values: * p<0.05, ** p<0.01, *** p<0.001

Source: 2010 Mail Response/Return Assessment data for PAHUE areas.

Table D2. Logistic Model of Mail Return Rate by Audience Segment

Odds ratios/[95% confidence intervals] Dependent var = Valid mail return by 4/19/10	**All Around Average**		**Economically Disadvantaged**		**Mobile Singles**		**Advantaged Owners**	
Log (GRP)	0.997		0.971		1.24	*	1.082	
	[0.944,1.053]		[0.877,1.076]		[1.025,1.500]		[0.984,1.189]	
Type of enumeration area								
Update/leave enumeration area	0.725	***	0.678	*			0.724	***
	[0.658,0.799]		[0.515,0.891]				[0.657,0.797]	
Military enumeration area	0.998		1.632	***	0.996			
	[0.806,1.235]		[1.534,1.735]		[0.814,1.219]			
Urban update/leave enumeration area	0.648	***	0.641	***			0.59	**
	[0.587,0.715]		[0.573,0.717]				[0.422,0.823]	
Outcome of 2000 collection at this address								
Address not in 2000 census	0.822	***	0.743	***	0.841		0.872	***
	[0.796,0.849]		[0.697,0.792]		[0.701,1.009]		[0.834,0.912]	
Enumerator completed return in 2000	0.634	***	0.639	***	0.797	***	0.591	***
	[0.621,0.648]		[0.622,0.658]		[0.769,0.826]		[0.572,0.610]	
Other outcome in 2000	0.719	***	0.697	***	0.754	**	0.719	***
	[0.689,0.749]		[0.654,0.743]		[0.633,0.899]		[0.662,0.781]	
Type of housing unit								
Multi-unit building	0.843	***	0.908	***	0.882	***	0.856	***
	[0.815,0.873]		[0.875,0.942]		[0.834,0.932]		[0.816,0.897]	
Trailer/mobile home/boat/RV, etc	0.743	***	0.764	***	0.774	**	0.704	***
	[0.718,0.769]		[0.738,0.792]		[0.663,0.904]		[0.673,0.737]	
Married couple	1.625	***	1.348	***	1.334	***	1.863	***
	[1.597,1.654]		[1.304,1.394]		[1.232,1.443]		[1.811,1.917]	
Household rents or occupies	0.632	***	0.693	***	0.534	***	0.556	***
without paying rent	[0.617,0.647]		[0.675,0.712]		[0.490,0.581]		[0.538,0.574]	
DMA pair for PAHUE								
Savannah/Columbia SC	1.191	*	1.324	*	1.367	**	1.34	
	[1.016,1.396]		[1.088,1.612]		[1.123,1.664]		[0.969,1.852]	
Flint-Saginaw/Toledo	1.445	*	1.184		1.51	**	1.534	**
	[1.020,2.046]		[0.814,1.723]		[1.161,1.963]		[1.163,2.022]	
Selma/Jackson MS	1.025		1.052		1.047		1.069	
	[0.941,1.117]		[1.000,1.108]		[0.940,1.167]		[0.908,1.259]	
Baton Rouge/Augusta	0.993		1.029		0.964		0.965	
	[0.813,1.213]		[0.838,1.263]		[0.895,1.038]		[0.783,1.188]	
Shreveport/Tallahassee	1.079	*	1.125	*	1.247	**	1.106	
	[1.019,1.143]		[1.005,1.260]		[1.082,1.436]		[0.883,1.387]	
Odessa-Midland/Lubbock	1.189		1.219		1.235		1.18	
	[0.869,1.626]		[0.966,1.538]		[0.835,1.828]		[0.711,1.960]	
Erie/Joplin-Pittsburg	1.301		1.153	***	1.282	***	1.395	
	[0.982,1.724]		[1.084,1.227]		[1.228,1.339]		[0.966,2.015]	

Table D2. Logistic Model of Mail Return Rate by Audience Segment (continued)

Odds ratios/[95% confidence intervals] Dependent var = Valid mail return by 4/19/10	**All Around Average**		**Economically Disadvantaged**		**Mobile Singles**		**Advantaged Owners**	
Household of unrelated people	0.709	***	0.715	***	0.709	***	0.716	***
	[0.680,0.740]		[0.686,0.745]		[0.630,0.798]		[0.690,0.744]	
Partnership activity scale (#1)	1		0.997		0.983		1	
	[0.993,1.007]		[0.989,1.006]		[0.965,1.002]		[0.989,1.011]	
Age of householder								
Householder age <=25	0.751	***	0.681	***	0.777	**	0.769	***
	[0.727,0.776]		[0.628,0.738]		[0.683,0.884]		[0.720,0.821]	
Householder age 41-64	1.579	***	1.721	***	1.603	***	1.552	***
	[1.537,1.623]		[1.676,1.769]		[1.534,1.675]		[1.497,1.609]	
Householder age 65+	3.006	***	2.945	***	3.141	***	3.125	***
	[2.858,3.162]		[2.763,3.138]		[2.862,3.446]		[2.907,3.358]	
Number of children age<18 in household								
1 child	0.963	**	1.017		0.88	***	0.939	***
	[0.943,0.983]		[0.984,1.050]		[0.853,0.907]		[0.928,0.950]	
2 children	1.015		1.02		0.916	*	1.013	
	[0.997,1.033]		[0.990,1.051]		[0.858,0.977]		[0.985,1.043]	
3-4 children	0.974	*	0.961		0.848	*	1.028	*
	[0.953,0.996]		[0.905,1.021]		[0.745,0.966]		[1.007,1.050]	
5-8 children	0.85	***	0.841	***	0.643	*	0.88	
	[0.798,0.905]		[0.790,0.896]		[0.466,0.887]		[0.698,1.111]	
9+ children	0.618		0.667		0.743		1.273	
	[0.303,1.262]		[0.436,1.019]		[0.200,2.762]		[0.751,2.159]	
Number of residents in household								
2 people	1.155	***	1.253	***	1.328	***	1.104	***
	[1.143,1.167]		[1.241,1.265]		[1.219,1.447]		[1.058,1.153]	
3 people	1.007		1.183	***	1.247	***	0.897	**
	[0.989,1.026]		[1.156,1.210]		[1.135,1.370]		[0.850,0.948]	
4 people	0.904	***	1.139	***	1.161		0.786	***
	[0.883,0.926]		[1.102,1.176]		[0.991,1.361]		[0.751,0.822]	
5 people	0.788	***	1.109	***	1.064		0.616	***
	[0.768,0.808]		[1.074,1.146]		[0.903,1.253]		[0.576,0.660]	
9-12 people	0.829	***	1.394	***	1.248		0.526	***
	[0.770,0.893]		[1.315,1.479]		[0.724,2.153]		[0.442,0.627]	
13+ people	0.293	***	0.421	*	0.139		0.271	**
	[0.169,0.507]		[0.211,0.840]		[0.006,3.327]		[0.132,0.553]	

Table D2. Logistic Model of Mail Return Rate by Audience Segment (continued)

Odds ratios/[confidence intervals] Dependent var = Valid mail return by 4/19/10	**All Around Average**		**Economically Disadvantaged**		**Mobile Singles**		**Advantaged owners**	
In area blanketed by replacement forms	1.256	***	1.277	***	1.23		1.427	***
	[1.143,1.381]		[1.198,1.362]		[0.955,1.584]		[1.257,1.621]	
In area where non-respondents received replacement form	1.157	***	1.216	***	1.106		1.168	**
	[1.101,1.216]		[1.113,1.329]		[0.904,1.354]		[1.067,1.278]	
Address is not residential	0.037	***	0.048	***			0.021	**
	[0.027,0.051]		[0.034,0.067]				[0.003,0.174]	
Received bilingual form (English/Spanish)	1.046	*	1.048		0.729	**	0.822	
	[1.006,1.088]		[0.980,1.122]		[0.625,0.850]		[0.504,1.342]	
Non-Hispanic Black	0.768	***	0.731	***	0.632	***	0.794	***
	[0.703,0.840]		[0.641,0.832]		[0.535,0.747]		[0.736,0.856]	
Hispanic	0.602	***	0.539	***	0.669	***	0.684	***
	[0.513,0.705]		[0.475,0.612]		[0.558,0.801]		[0.589,0.796]	

Note: Each column contains odds ratio with 95% confidence interval below it. Omitted categories: In mail-out/mail-back type of enumeration area, address had respondent return in 2000, owner, single-family house, Jacksonville/Little Rock DMA pair, single person household, non-Hispanic White, householder aged 26-40.

p-values: * p<0.05, ** p<0.01, *** p<0.001
Source: 2010 Mail Response/Return Assessment data for PAHUE areas.

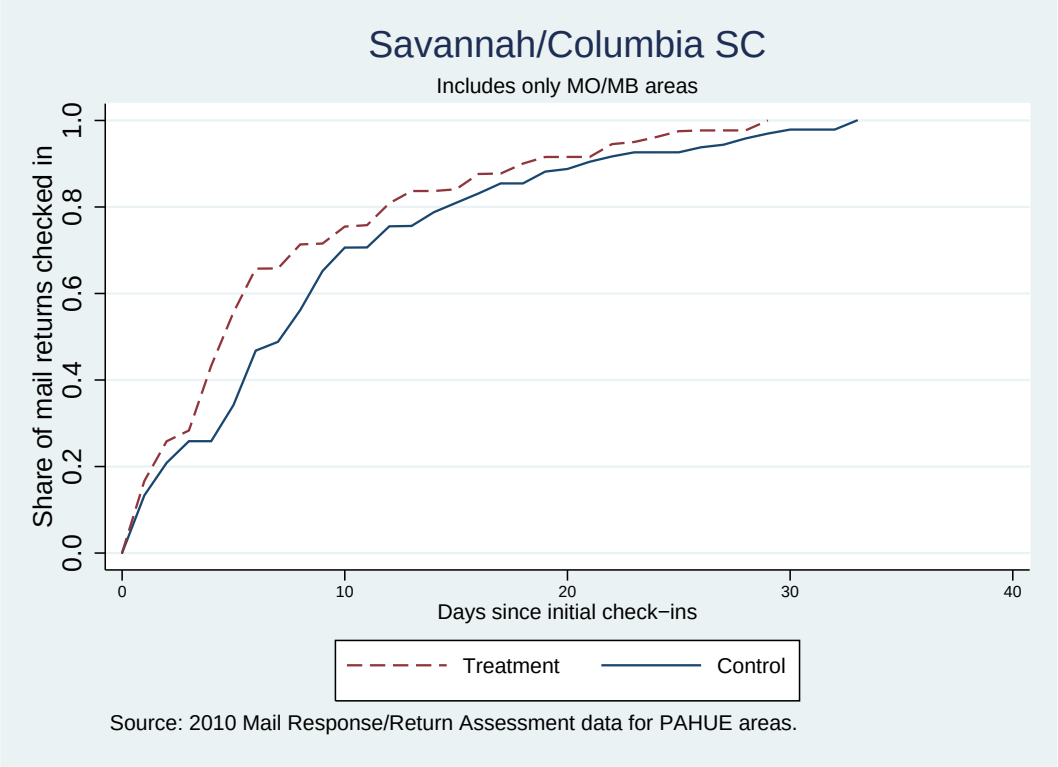
Appendix E: Additional Response Measures

Table E. Share of Returns by Response Type for PAHUE DMAs

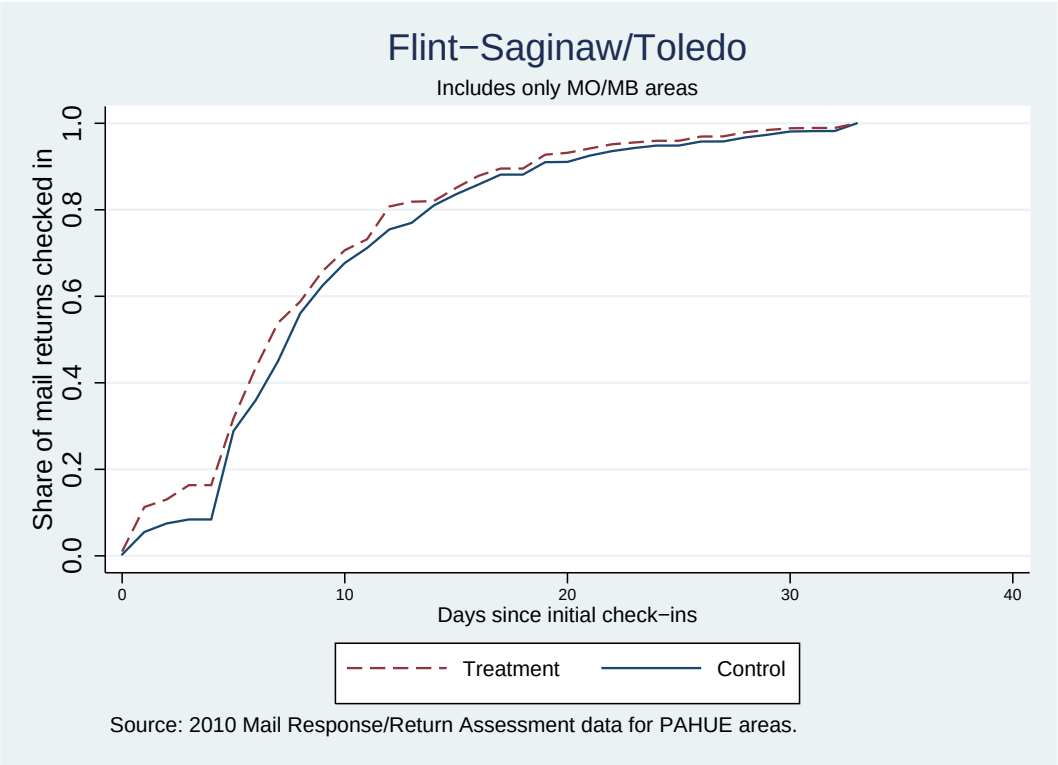
Pair	Treatment	DMA	Mail return by 4/19/2010	NRFU interview	Neighbor proxy	Mover proxy
Overall			75.1%	16.4%	6.3%	0.3%
Overall	Control		75.3%	16.7%	5.9%	0.3%
Overall	Treatment		74.9%	16.2%	6.6%	0.3%
1	Control	Columbia, SC	78.3%	14.6%	5.4%	0.2%
1	Treatment	Savannah	74.4%	16.9%	6.4%	0.3%
2	Control	Toledo	78.6%	13.9%	5.6%	0.2%
2	Treatment	Flint-Saginaw-Bay City	83.4%	10.8%	4.3%	0.2%
3	Control	Jackson, MS	72.4%	19.6%	5.9%	0.3%
3	Treatment	Montgomery (Selma)	73.1%	17.6%	7.3%	0.2%
4	Control	Augusta	76.0%	16.8%	5.1%	0.2%
4	Treatment	Baton Rouge	69.2%	19.8%	8.2%	0.3%
5	Control	Tallahassee-Thomasville	72.6%	17.7%	7.5%	0.3%
5	Treatment	Shreveport	72.4%	17.8%	7.2%	0.3%
6	Control	Lubbock	74.9%	16.7%	6.1%	0.3%
6	Treatment	Odessa-Midland	72.0%	18.5%	7.4%	0.3%
7	Control	Joplin-Pittsburg	76.3%	17.2%	5.0%	0.3%
7	Treatment	Erie	82.0%	11.5%	5.1%	0.2%
8	Control	Little Rock-Pine Bluff	73.1%	17.8%	6.5%	0.3%
8	Treatment	Jacksonville	72.8%	17.0%	7.2%	0.3%

Source: 2010 Mail Response/Return Assessment data for PAHUE areas.

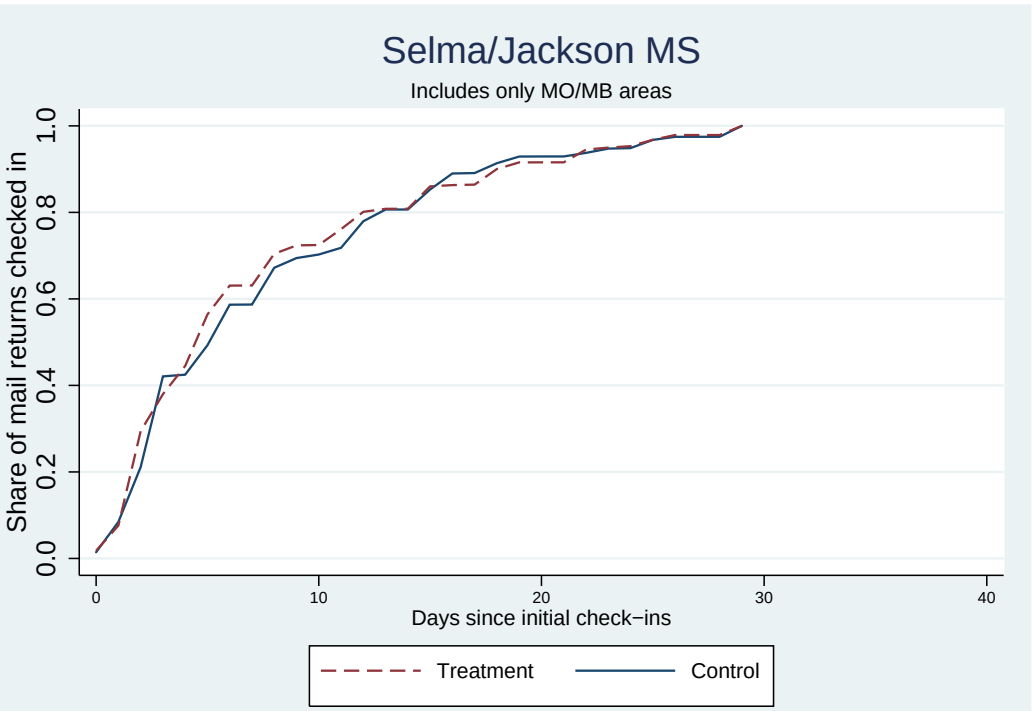
Appendix F: Timing Charts



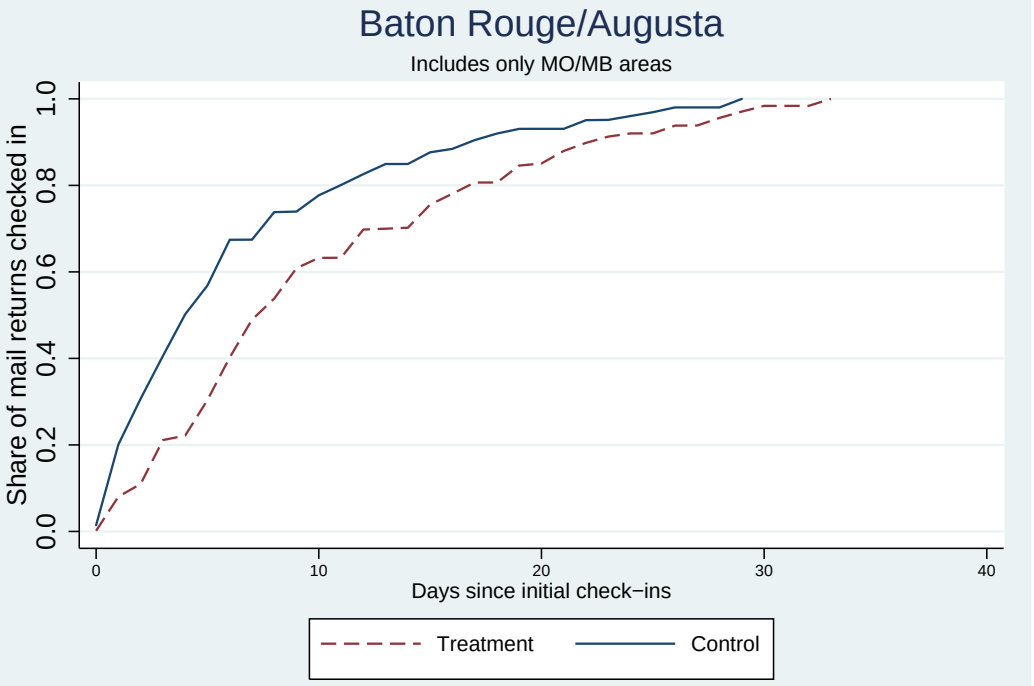
Source: 2010 Mail Response/Return Assessment data for PAHUE areas.



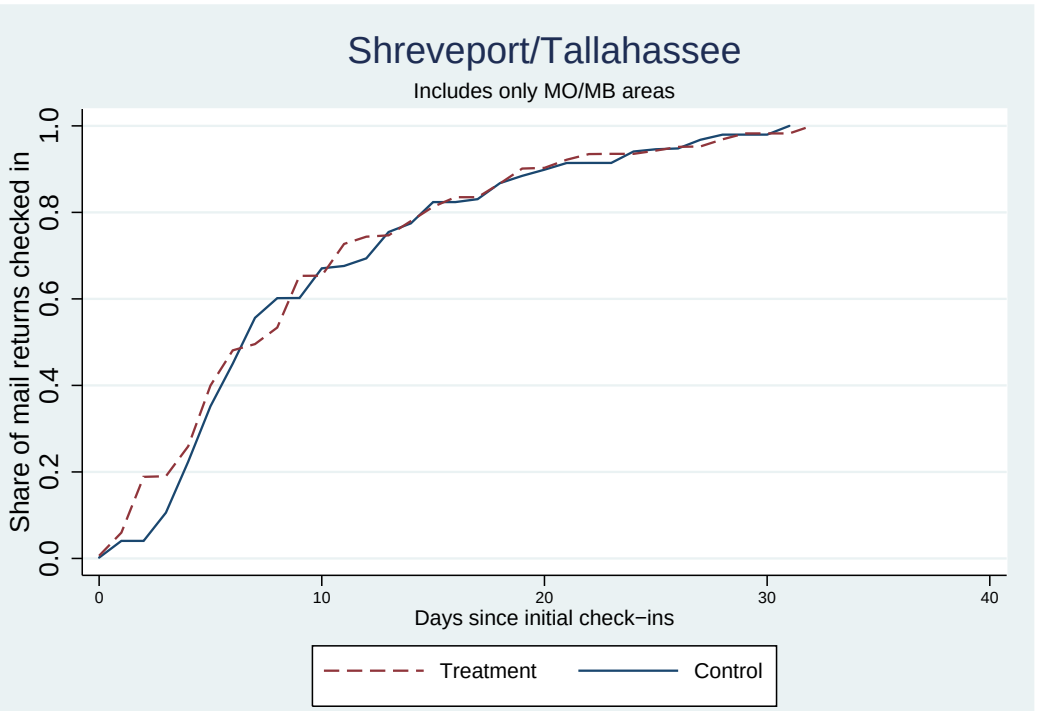
Source: 2010 Mail Response/Return Assessment data for PAHUE areas.



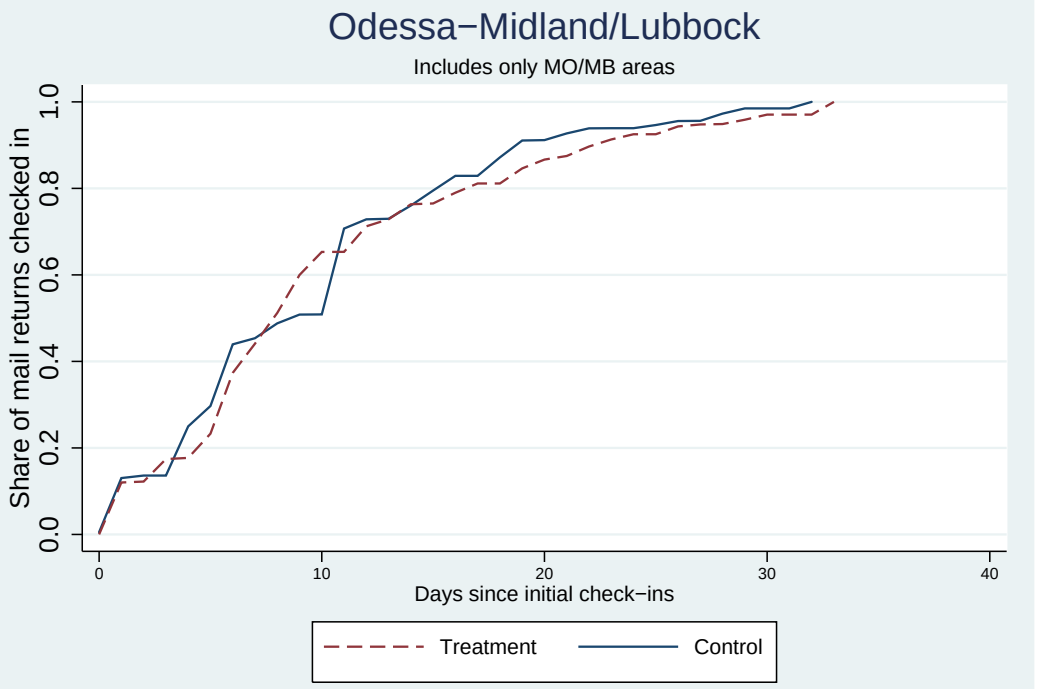
Source: 2010 Mail Response/Return Assessment data for PAHUE areas.



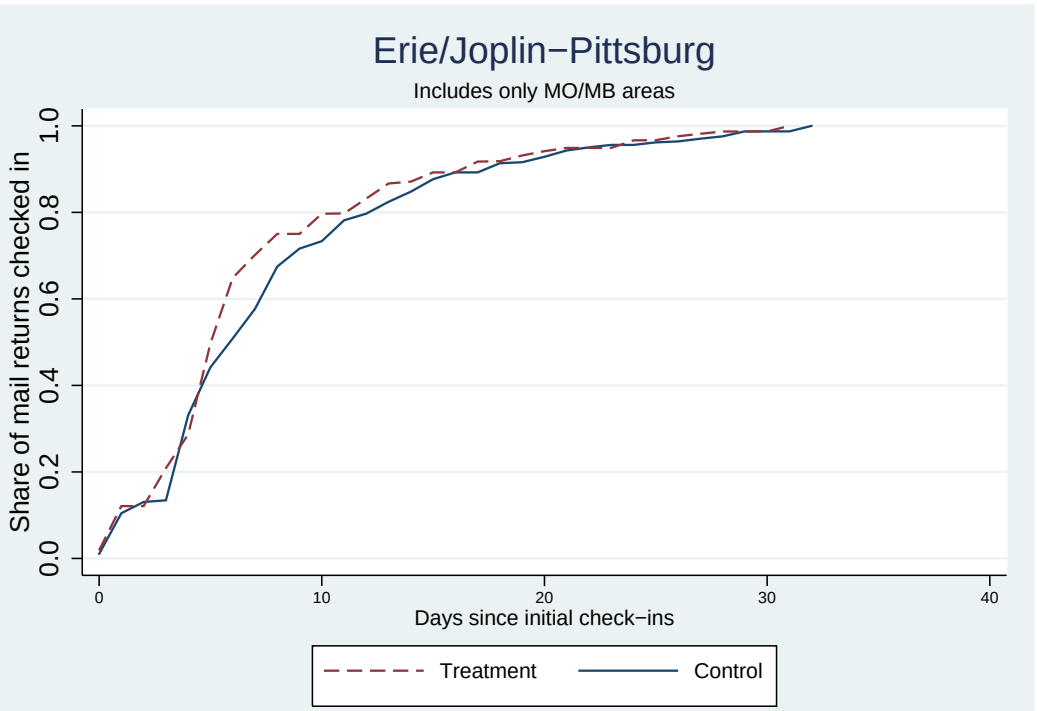
Source: 2010 Mail Response/Return Assessment data for PAHUE areas.



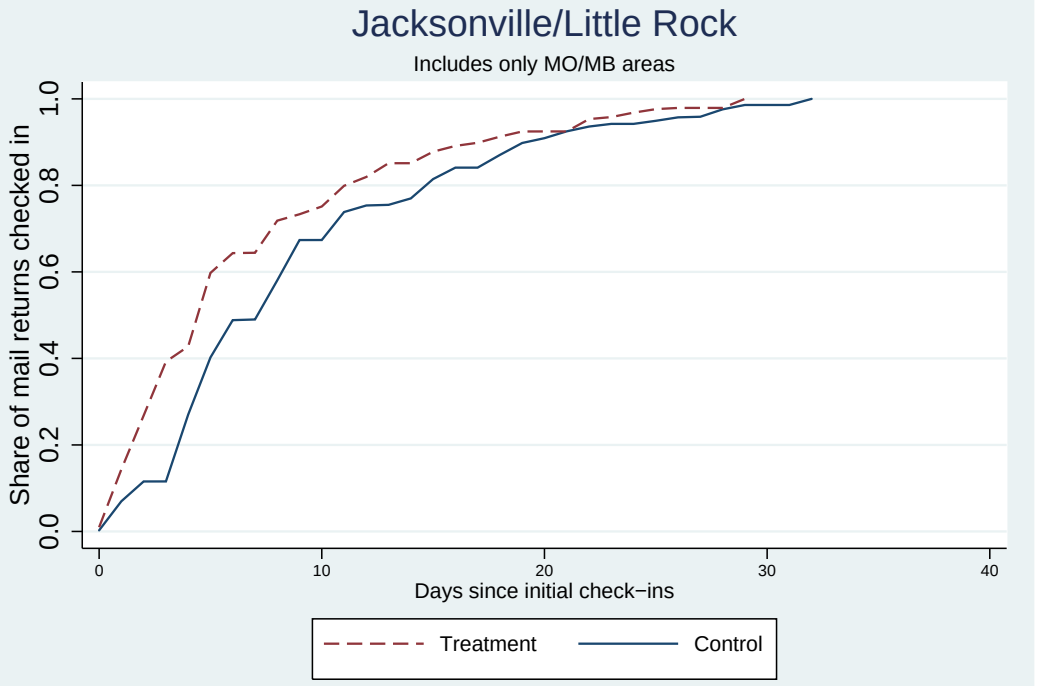
Source: 2010 Mail Response/Return Assessment data for PAHUE areas.



Source: 2010 Mail Response/Return Assessment data for PAHUE areas.



Source: 2010 Mail Response/Return Assessment data for PAHUE areas.



Source: 2010 Mail Response/Return Assessment data for PAHUE areas.