



This memorandum is intended for internal Census Bureau use and for external stakeholders who are interested in the Census Bureau's continuing research efforts. If you have any questions regarding the use or dissemination of this information, please contact Tasha Boone, Assistant Division Chief for Census 2010 Publicity Office (C2PO), ADCOM Directorate, at (301) 763-3977.

November 21, 2008

C2PO 2010 Census Integrated Communications Research Memoranda Series

No. 7

MEMORANDUM FOR      Distribution List

From:                      Kendall Johnson *[signed]*  
                                 Acting Chief, Census 2010 Publicity Office

Subject:                    Predicting Return on Investment Across Population  
                                 Segments in a Social Marketing Campaign

Contact Person:        Nancy Bates, C2PO Lead Researcher, (301) 763-5248

Attached is the C2PO 2010 Census Integrated Communications Research about Predicting Return on Investment Across Population Segments in a Social Marketing Campaign. The purpose of this research is to demonstrate social marketing return-on-investment (ROI) metrics and standards comparable to those used in product marketing. ROI is beneficial in providing evidence that there are measurable, positive outcomes of investing in a communication campaign and to guide and plan for the balance of communications across multiple platforms.

Attachment

## **Predicting Return on Investment across Population Segments in a Social Marketing Campaign**

Nancy Bates, U.S. Census Bureau  
Mary H. Mulry, U.S. Census Bureau<sup>1</sup>  
Tanya White, DraftFCB  
Darlene Billia, DraftFCB

### *Abstract*

For various reasons, social marketing campaigns have been slower to adopt many of the concepts and tools commonly used in the commercial marketing sector. In this paper, we demonstrate social marketing return-on-investment (ROI) metrics and standards comparable to those used in product marketing. The paper estimates ROI in a call-to-action social marketing campaign (participation in the 2010 Decennial Census). ROI is measured by the increase in the number of census forms completed and mailed back as a result of the campaign.

### *1. Introduction*

Census participation is critical to the public for several reasons. First, the decennial enumeration is the basis for the apportionment of seats in the House of Representatives (and is required in the Constitution). Population counts are also used for redistricting and the allocation of almost \$300 billion in federal funds each year. Second, for every 1% increase in mail response in the 2010 Census, roughly 75 million dollars will be saved in follow-up costs. Consequently, measuring return on investment (ROI) in a census is beneficial in two key ways: (1) to provide evidence that there are measurable, positive outcomes of investing in a communication campaign and (2) to guide and plan for the balance of communications across multiple platforms. Our analysis estimates ROI, meaning the yield in desired behavioral change for a given campaign expenditure, across different population segments prior to the launch of the 2010 Census communications campaign. This provides quantifiable benchmarks to help guide the allocation of limited resources as the communication campaign is being developed.

### *2. Social Marketing Campaigns and ROI*

Andreasen (1994) defines social marketing as the “application of concepts and tools from the commercial world to influence the voluntary behavior of target audiences to improve their lives and/or the society of which they are a part.” In the commercial world, methods of measuring product marketing ROI are product sales, market share, profits and the like. But unlike commercial marketers, social marketing often has difficult-to-measure goals tied to changes in behavior (e.g. stopping smoking, reducing teen pregnancies, or increasing the practice of “safe” sex).

---

<sup>1</sup> This report is released to inform interested parties of research and to encourage discussion. The views expressed are the authors’ and not necessarily those of the U.S. Census Bureau.

To measure the effectiveness of social marketing campaigns, researchers often turn to methods such as demographic surveys, hospital and health records, or official police and program statistics. These are often consulted before, during, and after the campaign to assess levels of behavior change. However, because the campaigns are not typically conducted as controlled experiments, many other factors can and do influence behavior. Consequently, it is very difficult to conclude causality as a result of social marketing campaigns. Even more difficult is finding a way to measure ROI.

An anti-smoking campaign might leverage several forms of media and marketing including television and print ads, billboards, merchandise, and grassroots advocacy groups. A survey might then be used to measure campaign awareness as well as smoking behavior and attitude changes. But using the survey data to distinguish ROI from television advertising expenditures versus community outreach is difficult, if not impossible, because controlling exposure to one versus the other is not possible. Nonetheless, that is precisely the challenge facing a once-in-a-decade census: How to best allocate limited advertising and community partnership resources to achieve the highest possible cooperation in the form of completing and mailing back census questionnaires?

### 3. *The 2010 Census*

The U.S. census is a massive undertaking – so large in fact it has been referred to as the “largest peacetime mobilization” in U.S. history (Hillygus, Nie, Prewitt and Pals, 2006 p. 2). The Census 2000 counted over 281 million people, and the projection for the 2010 population is over 308 million (U.S. Census Bureau 2007). The U.S. Constitution directs the census to count the resident population of the U.S. every 10 years for the apportionment of the U.S. House of Representatives. The emphasis on *resident population* means the census must include all who reside in the U.S. regardless of whether they are U.S. citizens, legal, or undocumented. The census also must count residents regardless of whether they speak or read English. Consequently, in contrast to many social marketing campaigns that focus on a single target population, the census encompasses and must reach *all* target populations. From a social marketing perspective then, the census is arguably one of the most comprehensive communication campaigns.

Census numbers are the data used to annually allocate billions of dollars in federal funds to states and small areas such as counties, places, and school districts. In the fund allocation formulas, often there is a fixed amount of money being allocated among the areas. These are known as ‘fixed pie’ allocations. For this type of allocation, the accuracy of an area’s share of the population is important in determining the amount of funds it receives and has very real consequences for the people residing within that area (Citro and Cohen 1985, Spencer 1980).

The methodology for taking the 2010 Census involves several steps. Most census questionnaires are delivered by mail although in some areas, U.S. Census Bureau employees deliver the questionnaires. The questionnaire asks for a list of those who live at the address as of April 1, 2010 in addition to a few other demographic questions. A member of the household fills out the form and returns it by mail. When a questionnaire

is not received, a census interviewer visits the address and conducts an interview in person.

The next census will occur April 1, 2010 and the U.S. Census Bureau has contracted with DraftFCB to develop a social marketing communication campaign. The campaign will consist of both paid advertising and a program of promotion and outreach to establish partnerships with a wide range of national and local organizations to implement promotional activities to educate the public about the importance of participating in the census. The 2010 integrated communication program has stated three goals:

- Increase the mail response rate<sup>2</sup>,
- Improve the overall accuracy and reduce the differential undercount, and
- Increase cooperation with enumerators.

The first two campaign goals can be competing. For example, allocating resources to increase the number of mail responses, may do little to decrease the differential undercount. However, resources allocated to decrease the differential undercount may have little impact on the goal of increasing overall mail response rate because the hard-to-count groups are a relatively small part of the population.

The mail response rate in Census 2000 was viewed as a success at 67 percent, much higher than the forecasted 61 percent (U.S. Census Bureau 2000). Some of this success was attributed to that fact that Census 2000 was the first to use paid advertising. The mail response bar for the 2010 Census has been set even higher at 69 percent. The 2010 Census has some improvements aimed at increasing response, such as the mailing of a replacement questionnaire to those who do not respond initially and switching to a short form only. Nevertheless, to achieve this goal, the campaign must identify areas and groups in the population that are likely to require targeting and special attention. To best plan and effectively balance the trade-offs of the competing census goals, it would be ideal to estimate the campaign ROI across population segments at the outset of the campaign.

#### *4. Segmentation and ROI in 2010 Census*

To allocate resources in the most efficient way, it is critical to begin a campaign by partitioning the population into segments since target audiences are seldom uniform in their perceptions and/or likely responses to marketing efforts. A segmentation scheme can guide the placement of paid media versus partnership activities and to some degree, the content of materials therein. Once the segmentation is defined, it is highly desirable to apply ROI estimates according to the size and characteristics of the segments. This allows us to calculate the expected increase in the number of desired outcomes (mail returns) for each segment .

---

<sup>2</sup> The mail response rate reflects the number of forms mailed back divided by the number mailed out (including both occupied and vacant households in the denominator).

The approach to the 2010 Census segmentation was hierarchical. First, we developed segments based on geo-demographic and social indicators. Second, we identified those segments that are hardest-to-count and isolated the factors that define them. Finally, we overlaid traditional marketing databases containing media consumption data to inform where and how to buy media for the different segments.

By the very nature of the decennial census, the U.S. Census Bureau maintains a vast quantity of data about who participates in the census and by what method. Consequently, unlike some call-to-action campaigns, the U.S. Census Bureau has historical data related to the behavioral outcome of interest (mailback behavior from previous censuses). Two such sources were used to estimate ROI. First is the U.S. Census Bureau Census 2000 Planning Database (PDB). This is a tract-level database containing a range of housing, demographic, and socioeconomic variables correlated with mail response (Bruce and Robinson 2006). The database is publicly available on the U.S. Census Bureau website (<http://www.census.gov/procur/www/2010communications/library.html>). The PDB was then merged with an operational file from Census 2000 that contained mail response rates across tracts. This yielded a macro-level indicator of behavior.

The second data source consisted of a household-based multi-wave sample survey conducted in conjunction with Census 2000. The National Opinion Research Center conducted the survey, called the Partnership and Marketing Program (PMP) survey. The PMP survey measured census awareness, attitudes, and behavior over the life of the Census 2000 campaign. It also measured exposure to a variety of communication platforms including mass media such as radio, television, newspapers, as well as community-based communications including schools, religious groups, local officials, and community meetings. Additionally, for two waves of survey data, respondents were linked back to official census databases to determine if the household mailed back a census form (see Wolter et al, 2002). This provided a micro-level indicator of behavior and together, the PDB and PMP survey yielded an estimate of ROI across different population segments.

Finally, to efficiently allocate resources across segments, we tapped a third data source reflecting more traditional consumer survey databases. We overlaid two commercial consumer surveys (Mediamark Research Institute and Simmons) to further define the segments and guide where to place communications across different platforms. Examples might include where to emphasize the grass-roots partnership efforts over paid advertising, where to place in-language media, and what media channels to tap for which segments.

##### *5. What Did We Find?*

Our approach was to distinguish among groups in the population that are difficult to count from those that are easier to count. Using the PDB and cluster analysis, we created eight sociodemographic clusters based upon population tract characteristics that are correlated with the likelihood of mailing back a census form. See Bates and Mulry (2007) for detailed methodology. Specifically, we concentrated on 12 variables used to

comprise a “hard to count” score in previous censuses (Robinson, Johanson, and Bruce 2007). These included:

- % vacant units,
- % non-single family attached/detached units;
- % renter occupied units;
- % units with >1.5 persons per room ;
- % non-spousal units;
- % units without phone;
- % people below poverty level;
- % units receiving public assistance;
- % people unemployed;
- % linguistically isolated households,
- % moved within last year, and
- % without a high school degree.

Although mail response was not used directly in forming the clusters, Table 1 clearly illustrates that the clusters capture where the low, medium, and high mail response tracts were located.

**Table 1. Mail Return Rate, Number of Tracts, and Occupied Housing Units by Cluster**

#	Cluster Name	2000 Mail Return Rate	Total Occupied Housing Units		Number of Tracts
			Number (in millions)	Percent	
1	All around average I (homeowner skewed)	77.3%	36.5	35%	21,174
2	All around average II (renter skewed)	74.2%	16.5	16%	8,957
3	Econ. Disadvantaged I (homeowner skewed)	66.5%	6.6	6%	5,230
4	Econ. Disadvantaged II (renter skewed)	58.0%	3.0	3%	2,574
5	Ethnic Enclave I (homeowner skewed)	69.8%	3.4	3%	2,440
6	Ethnic Enclave II (renter skewed)	63.6%	2.5	2%	1,754
7	Young/mobile/singles	67.1%	8.0	8%	4,073
8	Advantaged Homeowners	83.2%	26.8	26%	16,506

Note: The mail return rate is the percentage of occupied housing units that returned a mail form

#### *6. Description of clusters*

The cluster analysis revealed eight distinct groups each with varying levels of mail return behavior in 2000 and each with unique demographic, housing, and socioeconomic characteristics. In three instances, pairs of clusters appear closely related to one another with homeownership/renter status as the distinguishing feature (i.e., All Around Average I and II; Economically Disadvantaged I and II; and Ethnic Enclave I and II).

### Cluster 1: All Around Average I (homeowner skewed)

This group has the largest number of occupied housing units and had the second highest mail return rate in 2000<sup>3</sup>. They are best described as the “average Joe” cluster in that they are close to average on every one of the hard-to-count variables. Around 28% of the housing units are not single-family structures, only one-quarter are renters, and slightly less than half (45%) are in non-spousal households.

Unemployment, poverty, education and mobility levels are all close to national averages. The tracts are fairly representative of the national average racial breakouts but have above-average percentage of non-Hispanic whites (80%) slightly below-average blacks (9%), 2% Asian or Native Hawaiian/Pacific Islander (NHPI) and 1% American Indian/Alaska Native (AIAN). Tracts in this cluster contain about 7% Hispanics which is below the national average. Around one-quarter of the population is under age 18 and about 15% are over 65. This group is the largest cluster representing about 35.5 million occupied housing units (about 35% of the total). This cluster has the largest percentage of rural tracts<sup>4</sup> (on average around 37% are rural).

### Cluster 2 - All Around Average II (renter skewed)

This cluster is also somewhat unremarkable and “average” on most of the hard-to-count variables. About the only distinguishing characteristic is an above average number of households renting and in multi-units. This group of tracts is slightly more racially diverse than Cluster 1 (12% black, 11% Hispanic, and 69% non-Hispanic white) and is also much more urban and densely populated. However, like Cluster 1, this group is relatively large (represents around 16% of all occupied housing units).

### Cluster 3 – Economically Disadvantaged I (homeowner skewed)

This cluster reflects households that are economically disadvantaged, but not as much as Cluster 4. One noticeable difference is that this cluster has fewer renters than Cluster 4 (less than half rent – 46%). Nonetheless, these tracts have a high percentage in poverty, without a high school education, and on public assistance. Above average unemployment is also characteristic of this cluster. Blacks comprise about one-half (49%) of the population in these tracts – the second largest black population next to Cluster 4. This cluster has above-average number of children (29% are younger than 18). This group represents about 6% of the total occupied housing units. The overwhelming majority of tracts in this cluster are urban (92% urban on average).

---

<sup>3</sup> The mail return rate is the percentage of occupied housing units to which the U.S. Census Bureau mailed a census form that return it. Vacant housing units are not included in the denominator.

<sup>4</sup> “Urban” is defined as housing units located within urbanized areas (UAs) or urban clusters (UCs). A UA consists of areas containing 50,000 or more people while a UC consists of areas with a least 2,500 people but fewer than 50,000. “Rural” consists of areas located outside of UAs and UCs (U.S. Census Bureau, 2001).

#### Cluster 4 – Economically disadvantaged II (renter skewed)

This cluster had the lowest mail return rate of any group and also the highest hard-to-count score. Close to three-quarters of the households in these tracts contain non-spousal renters in multi-units (especially 10+ units). These tracts also have the highest poverty, public assistance, and unemployment than any other cluster. This cluster most closely resembles Cluster 3 but has far fewer homeowners (on average, 81% of households are renters). Like Cluster 3, this group contains a higher-than-average percentage of Blacks (54%) but also has an above-average percentage of Hispanics (21%). This cluster reflects the most urban of all clusters (99.9% urban on average). This cluster represents about 3% of the total occupied housing units.

#### Cluster 5 – Ethnic enclave I (homeowner skewed)

This cluster is characterized by above-average crowding and poverty, public assistance, unemployment and low education. However it also contains a *below-average* percentage of non-spousal households and above-average percentage of children. It looks most like Cluster 6 with the following differences: lower occurrence of linguistic isolation, lower mobility, higher homeownership, and fewer Asians. This cluster is also less urban and less densely populated than Cluster 6. This group is predominantly Hispanic (61%) with 24% non-Hispanic white, 8% black, and 5% Asian or NHPI.

#### Cluster 6 – Ethnic enclave II (renter skewed)

This cluster has the second-highest HTC score of any cluster (and second lowest mailback rate). This cluster has above-average presence of children and is characterized by multi-unit structures with at least 10 units. This group is exclusively urban, the most densely populated of clusters, and characterized by crowded housing. On average, half of persons residing within this cluster lack high school degrees. These tracts are predominantly comprised of Hispanics (59%) and Asians (11%) with only 19% non-Hispanic white, 9% black, and 1% AIAN.

This cluster contains tracts with high levels of linguistic isolation (on average, around 31%). In some tracts, this ranges as high as 79% of households where Spanish is spoken at home or no household member 14 or older speaks English very well. Likewise, other tracts have as high as 74% of households where an Asian/Pacific Islander language is spoken at home or no household member over 14 speaks English very well. This group is overwhelmingly renters (75%). It also has high rates of poverty, unemployment, and public assistance. This is the smallest of the 8 clusters, representing only 2% of the total occupied housing units. As such, increases to response rates will yield a smaller number of actual mail forms compared to the other clusters.

#### Cluster 7 – Young/single /mobiles

This cluster had a similar mail return rate and hard-to-count score as Cluster 3 but looks very different. The overwhelming majority of households are non-spousal renters located



in multi-units (especially structures with at least 10 units). The people in these tracts have higher than average education along with very high mobility. The tracts are densely populated and almost exclusively urban. These tracts have a below average percentage of children (17%). This cluster has a relatively high percent of group quarters (4%), possibly reflecting college campuses. These tracts probably include younger singles in school or just out of school and into the workforce for the first time.

This cluster is racially diverse with above-average percentage Asian (7%) and the majority non-Hispanic white (59%) followed by black (17%). This group represents about 8% of the total occupied housing units.

#### Cluster 8 – Advantaged homeowners

The tracts in Cluster 8 had the highest mail back rate and the lowest HTC score in 2000. As such, these tracts have a very low percentage of renters, few multi-units structures, very low levels of poverty and unemployment, low mobility, and few non-spousal households. This cluster is indicative of stable homeowners who reside in spousal-households in single-unit houses, about one-quarter of which are located in non-urban areas. This group of tracts is the least racially diverse of all clusters with 85% non-Hispanic white and only 4% black, 5% Hispanic, 4% Asian or NHPI and less than 1% AIAN. It is also the least densely populated cluster as measured by population per square mile. This group is the second largest behind Cluster 1 reflecting 26% of the total occupied housing units.

#### 7. *Predicting ROI*

Once the clusters were defined, we next used logistic regression models fit with data from the 2000 PMP survey that estimated the probability of mailing back a census form under different levels of reported campaign exposure. We estimated the probability of returning a form for those who had not seen any of the communications and assumed the probability with no awareness was equal to the probability if there were no communications campaign (see Mulry and Keller, 2007 for detailed methodology). Next we estimated the probability for those who have “average” scores on communication exposure. The difference in the probabilities was our estimate of the increase in the probability of response associated with the campaign, the estimated ROI. Then multiplying the difference in the probabilities by the population size produced an estimate of the number of additional returns. With this approach, we estimated the ROI and the additional number of forms for the nation as whole and within each population segment.

Overall, the approach estimates an increase of 5.5 million mail returns in all eight clusters combined associated with the Census 2000 communications campaign. Since there are 102.8 million occupied housing units in the mailout/mailback areas in the eight clusters, this implies that the advertising campaign and partnership program increased the mail return rate by about 5.3 percentage points in these areas.

These results are corroborated somewhat by the increase in mail response observed by a nationwide test of the American Community Survey (ACS). During the months January through March, the mail response was 5 to 9 percentage points higher in 2000 than in those same months in 2001 when there was no advertising by the Census Bureau (Bentley, Trancroto, and Hill 2006).

Tables 2 and 3 show how the results of the logistic regression models may be used to classify the clusters by high, medium, and low for the increase in *probability* of response associated with advertising and marketing, and for the increase in *number* of responses associated with advertising and marketing.

**Table 2. Clusters classified by estimated percentage increase in Census 2000 mail return rate associated with PMP.**

Very high ( $\geq 8\%$ )	High ( $< 8\%$ and $\geq 7\%$ )	Medium ( $< 7\%$ and $\geq 5.3\%$ )	Low ( $< 5.3\%$ )
Cluster 4 (Econ. Disadv. II)	Cluster 3 (Econ. Disadv. I)	Cluster 7 (Young/mobile/single)	Cluster 1 (All avg. I)
		Cluster 2 (All avg. II)	Cluster 8 (Advan. Homeowners)
		Cluster 5 (Ethnic enclave I)	
		Cluster 6 (Ethnic enclave II)	

**Table 3. Clusters classified by estimated increase in numbers of occupied housing units mailing back a Census 2000 form associated with PMP.**

Very high ( $\geq 1.5$ mil)	High ( $< 1.2$ mil and $\geq 800,000$ )	Medium ( $< 550,000$ and $\geq 400,000$ )	Low ( $< 300,000$ )
Cluster 1 (All avg. I)	Cluster 2 (All avg. II)	Cluster 3 (Econ. Disadv. I)	Cluster 6 (Ethnic enclave II)
	Cluster 8 (Advan. Homeowners)	Cluster 7 (Young/mobile/single)	Cluster 4 (Econ. Disadv. II)
			Cluster 5 (Ethnic enclave I)

The clusters with the largest increase in probability of mail return associated with the campaign were not necessarily the clusters that produced the largest increase in number of mail returns. For the larger groups a small increase in the probability of mailing back a census form produces a large number of forms. For example, the estimated percentage increase for the Advantage Homeowner cluster with almost 27 million occupied households is the lowest at 4.1 percent, but the estimated increase in number of mail returns is over one million.

## 8. What Are the Implications?

The estimated ROI and segmentation clusters help address strategic decisions related to the three census goals. As Tables 2 and 3 illustrate, the ROI varies by cluster but also by which of the three census goals is in question. The ROI estimates coupled with the size and characteristics of the clusters helps solve the potential paradox of decreasing the differential undercount while increasing the number of mail returns. Once we know the value for both, we can calculate the expected increase in the *number* of mail returns for each cluster as opposed to the *percentage* increase. For example, our Census 2000 model indicates that the highest percentage increase in mail return associated with the social marketing campaign is for the most economically disadvantaged cluster (over 8 percent increase). However, in terms of absolute number of mail returns, the biggest yield comes from the cluster with the largest population (yet had the lowest estimated percentage point increase associated with campaign at 5%).

Additionally, the patterns of media usage vary across clusters even though communication modes were not considered in their formation. We discovered this when we examined usage patterns of key media channels within clusters based on results from merges between the PDB clusters and the two consumer surveys (Simmons and MRI).<sup>5</sup> We found that tracts in both of the All Around Average clusters were moderate users of magazines, newspapers, and radio but that All Around Average I tracts had considerably lower Internet use compared to All Around Average II. Tracts in both clusters may not require much targeting beyond the broad campaign elements designed to hit all sectors of the population. Because both clusters represent just over half of all occupied housing units, even modest percentage gains in response will translate into a large absolute number of mail returns (and thus cost savings from large reductions in personal visit follow-ups).

The two economically disadvantaged clusters (clusters 3 and 4) will likely require targeting and special attention to achieve the desired mailback cooperation. Tracts in the Economically Disadvantaged I cluster are probably good candidates for word-of-mouth (WOM) activities of the partnership campaign (e.g. promoting census participation through trusted community leaders) since this segment exhibited low usage of media with the exception of cable TV. The Economically Disadvantaged II cluster was among the heaviest users of all media types with the exception of the Internet. Because these clusters contain above average percentage of minorities, a campaign that produced a high mail return rate in these tracts could help decrease the differential undercount.

With the exception of radio listening, the Ethnic Enclaves I and II do not appear very reachable by traditional English-language media channels or by the Internet. Both clusters scored very high as non-users of mainstream media including magazines, newspapers, and out of home advertising. However, both clusters reported high Spanish-language media usage. Consequently, both are good candidates for in-language materials delivered by in-language media or WOM delivered by community-based organizations and local leaders. For the tracts with a high percentage of linguistic isolation and where

---

<sup>5</sup> See technical appendix for data merge methodology.

Spanish is spoken, mailing bi-lingual census forms in English and Spanish may aid in increasing the mail return rate. Improving response to the census in these clusters could help decrease the differential undercount.

The consumer survey data for the young, mobile singles (Cluster 7) indicated very high usage of the Internet and above average magazine and newspaper readership. Because of the heavy on-line presence, this cluster is a good candidate for conveying messages through web-based social networks and media. Since the mobility of persons in this cluster makes them harder to count, increasing their mailback rate would improve the census count and reduce followup costs.

As with the young, mobile single cluster, the Advantaged Homeowners tend to also be high Internet users. This cluster also showed high usage of other mainstream media including TV, radio, and print readership. Mass media will probably work well to reach this cluster and because this segment contains 26 percent of U.S. households, even a modest increase in the mailback rate will result in a substantial reduction in followup costs.

#### *9. Future Research*

Many challenges face the U.S. Census Bureau and DraftFCB as we move ahead with the 2010 Census Integrated Communication Campaign. While the ROI metrics described here serve as a tool for planning the 2010 Census campaign, several questions remain. First is the problem of disaggregating ROI for the paid media versus community-based partnership activities. In the PMP survey, the self-reported exposure measures were simply too correlated to cleanly discern between them. Consequently, our ROI reflects the estimated impact of the entire 2000 communications campaign without any idea how much was the result of paid media versus local grass-roots efforts. Second, the ROI estimates are based on survey recall questions of census advertising exposure without the aid of visuals or audios. This casts further doubt on the precision of the ROI estimates and how much we should rely upon them when developing budget allocation models.

Finally, the data in the PDB is now seven years old. Tract characteristics can and do change over the decade, making the geographic location of clusters less accurate. To address whether the core composition of the clusters has shifted significantly over time, we are analyzing another rich U.S. Census Bureau datasource, the American Community Survey (ACS). The ACS is a large nationwide household survey conducted monthly that collects all the items used to calculate the Census 2000 hard-to-count score. We will compare the cluster distributions for the hard-to-count variables in the 2006 ACS to those from the 2000 PDB. Since we are examining clusters, not individual tracts, we do not expect to find dramatic changes, but if we do, we will adjust the segmentation scheme accordingly.

### *Acknowledgements*

The authors wish to acknowledge Joe Panzarella and Peter Luo of DraftFCB and John Mossawir of U.S. Initiative for performing the data analysis of the MRI and Simmons consumer survey databases merged with the 2000 Enhanced Planning Database. We also wish to thank Marie Pees of the U.S. Census Bureau for providing additional data files and variables that made the ROI analysis possible.

### *References*

- Andreasen, Alan R. (1994) "Social Marketing: Definition and Domain". *Journal of Public Policy & Marketing*. 13(1). American Marketing Association. Chicago, IL. 108–114.
- Bates, N. and Mulry, M. H. (2007) "Segmenting the Population for the 2010 Census Integrated Communications Program". Unpublished paper dated October 22, 2007. Statistical Research Division. U.S. Census Bureau.
- Bentley, M., Tancreto, J., and Hill, J. (2006) "Research on Census Response Rates and Partnerships and Marketing". Internal report dated September 6, 2006. Decennial Statistical Studies Division, U.S. Census Bureau, Washington, DC.
- Borsa, F. R. and Hough, C. L. (2004) "Data Collection in Census 2000". Census 2000 Testing, Experimentation, and Evaluation Program. Topic Report No. 13, TR-13. U.S. Census Bureau. Washington, DC. <http://www.census.gov/pred/www/rpts/TR13.pdf>
- Bruce, A. and Robinson, J.G. (2006). "Tract- Level Planning Database with Census 2000 Data." U.S. Department of Commerce, U.S. Census Bureau: Washington, D.C. <http://www.census.gov/procur/www/2010communications/library.html>
- Citro, C. F. and Cohen, M. L. (1985) *The Bicentennial Census: New Directions for Methodology in 1990*. National Academy of Science Press. Washington, DC.
- Hillygus, D.S., Nie N., Prewitt K. and Pals, H. (2006). *The Hard Count: The Political and Social Challenges of Census Mobilization*. The Russell Sage Foundation: New York: New York.
- Mulry, M. H. and Keller, J. K. (2007) "Investigation of the Relationship between Census Mail Response and the Marketing and Partnership Campaign". *2007 Joint Statistical Meeting Proceedings, Section on Survey Research Methods*. [CD-ROM]. American Statistical Association. Alexandria, VA. 3856-3863.
- Robinson, J.G., Johanson, C. and Bruce, A. (2007). "The Planning Database: Decennial Data for Historical, Real-time, and Prospective Analysis". Paper presented at the 2007

Joint Statistical Meetings, Salt Lake City, Utah. Population Division, U.S. Census Bureau, Washington, Dc.

Spencer, B. D. (1980) *Benefit-Cost Analysis of Data Used to Allocate Funds*. Springer. New York.

U.S. Census Bureau (2007) *Statistical Abstract of the United States: 2008 (127th Edition)*. U.S. Census Bureau. Washington, DC. <http://www.census.gov/statab/www/>

U.S. Census Bureau (2000) “‘Well Done, America!’ Nation Achieves 67 Percent Response Rate in Census 2000, Two Points Higher Than 1990”. Press Release dated September 19, 2000. U. S. Census Bureau. Washington, DC. [http://www.census.gov/Press-Release/www/releases/archives/census\\_2000/000657.html](http://www.census.gov/Press-Release/www/releases/archives/census_2000/000657.html)

Wolter, K. et al (2003) “Partnership and Marketing Program Evaluation”. Census 2000 Evaluation D.1. Prepared by NORC. U.S. Census Bureau. Washington, DC. <http://www.census.gov/pred/www/rpts/D.1.pdf>

Technical Appendix

***Mediamark Research Institute consumer survey merge with the U.S. Census Bureau 2000 Planning Data Base (PDB)***

Mediamark conducts an ongoing study of the adult population of the United States. This study, conducted continuously since 1979, surveys the demographics, product usage, and media exposure of all persons aged 18 and over in the contiguous 48 states. The sample size is approximate 26,000 people per year. DraftFCB analyzed the 2007 MRI Doublebase by using a combination of the two previous spring studies (spring 2006 and spring 2007).

In total, 51,182 respondents were included. Each of the 51,182 responders was geocoded and linked to the PDB by GIDtract then assigned the cluster code of the tract. 948 MRI respondents were not matched with the U.S. Census Bureau 2000 Planning Data Base (PDB) because it excludes “nonrepresentative” tracts that contain less than 100 housing units and/or 250 people, or that have >50% population in group quarters or that have >35% population over age 65.

The following table provides the distribution of the 2007 MRI Doublebase tract responders by cluster. For comparison, the tract count of the PDB is also listed.

Cluster	PDB Tracts	PDB %	PDB Cumul %	MRI Tracts %	MRI Cumul %
Average I Homeowner Skew	21,174	33.8%	33.8%	28.8%	28.8%
Average II Renter Skew	8,957	14.3%	48.0%	15.6%	44.3%
Econ Disad I Homeowner Skew	5,230	8.3%	56.4%	3.4%	47.8%
Econ Disad II Renter Skew	2,574	4.1%	60.5%	1.8%	49.5%
Ethnic Enclave I Hmowner Skew	2,440	3.9%	64.4%	1.9%	51.4%
Ethnic Enclave II Renter Skew	1,754	2.8%	67.2%	1.9%	53.3%
Young / Mobile / single	4,073	6.5%	73.7%	8.3%	61.7%
Advantaged Homeowners	16,506	26.3%	100.0%	38.3%	100.0%
<b>Total</b>	<b>62,708</b>				

***Simmons national consumer survey merge with the U.S. Census Bureau 2000 Planning Data Base (PDB)***

Simmons Research, a part of Experian, has chronicled the American consumer for over 50 years. DraftFCB conducted a data merge of the Simmons Spring 2007 one year and two year samples with the U.S. Census Bureau 2000 Planning Data Base (PDB). The linkage was conducted by acquiring the census tract definition of each Simmons respondent and assigning the PDB census cluster designation. Of 49,813 respondents over the two year sample, 49,023 were successfully appended with a cluster designation. The one year sample was used as the base survey for most of the media analysis. The two year sample was used for analyses where there were sample size challenges (i.e, subsamples within African American universe).

Overall the unweighted dispersion of Simmons respondents by census tract across clusters were 95% correlated with the dispersion of census tracts within the PDB.

Cluster	PDB Tracts	PDB %	PDB Cumul %	Simmons Tracts %	Simmons Cumul %
Average I Homeowner Skew	21,174	33.8%	33.8%	30.0%	30.0%
Average II Renter Skew	8,957	14.3%	48.0%	14.1%	44.1%
Econ Disad I Homeowner Skew	5,230	8.3%	56.4%	3.9%	48.0%
Econ Disad II Renter Skew	2,574	4.1%	60.5%	2.6%	50.6%
Ethnic Enclave I Hmowner Skew	2,440	3.9%	64.4%	5.8%	56.4%
Ethnic Enclave II Renter Skew	1,754	2.8%	67.2%	4.5%	60.9%
Young / Mobile / single	4,073	6.5%	73.7%	5.7%	66.6%
Advantaged Homeowners	16,506	26.3%	100.0%	33.5%	100%
<b>Total</b>	<b>62,708</b>				