

Non-response in a National RDD Survey:

Analysis of Basic Effort and Extra Efforts

By Evans Witt and Jonathan Best

Princeton Survey Research Associates International

Executive Summary

In the past two decades, the research profession has faced increasing levels of potential respondents to surveys failing to complete interviews. The levels of such non-response raise questions about the representativeness and the validity of surveys and the data they provide.

The U.S. Office of Management and Budget asked the Pew Internet & American Life Project and the University of Illinois, working with a grant from the U.S. Institute of Museum and Library Services¹, to conduct additional analysis and an experiment on non-response as part of a larger national survey in 2007 exploring the role of libraries in the internet age. The additional analysis and experiment had three elements.

First, the interviewing effort on a sample of the telephone numbers was doubled. Where the design called for a maximum of 10 calls to each number, at least 1,500 numbers were called at least 20 times. The results of that extra effort were compared with the results of the standard 10-call effort.

Second, an analysis was conducted of survey results from the base 10-call design by segmenting interviews by the amount of effort actually required to get results.

Third, the total sample of telephone numbers and the subset that provided completed interviews using the 10-call design were analyzed to determine what kinds of communities are under- and over-represented in completed interviews.

The results of the analysis include:

- ✓ Doubling the interviewing effort to 20 calls produced 84 additional interviews. The results from these extra-effort interviews varied only occasionally and marginally from results in the base study.
- ✓ The extra interviewing effort had the expected impact: that is, those who are harder to reach in general were those reached with extra effort. Younger adults, working adults and those with college degrees, thus, were a larger share of the extra-effort completes.
- ✓ Analyzing the base survey results by the level of effort required to achieve an interview found few statistically significant differences.

¹ OMB Clearance Number, 3137-0070, expiration date 06.30.2010.

- ✓ Analyzing the complete RDD sample and the completed interviews by community characteristics showed that interviews are hardest to complete in urban areas and easiest to complete in rural areas. While there appear to be no significant variations across communities by average household income, areas with higher minority populations (both Hispanic and African-American) were less productive in terms of interviewing, paralleling the finding on urban areas.

Introduction

In 2007, the Pew Internet & American Life Project and the University of Illinois partnered with funding from the U.S. Institute of Museum and Library Services² to conduct a national survey to explore the role of libraries, the internet and other avenues in how Americans seek information and assistance on matters often related to the government.

The U.S. Office of Management and Budget, in approving this project as part of the normal federal process, asked for an experiment to look at some of the issues surrounding increasing non-response to telephone surveys of the general American population. In specific, OMB approved an experiment that called for additional effort to seek completed interviews from telephone numbers who did not provide interviews in the normal course of the survey effort and for added analysis of the basic data.

This is a report on that experiment.

Non-Response

In the past two decades, survey professionals have faced with increasing levels of potential respondents to a survey failing to complete interviews. These failures to complete interviews result from a variety of factors, but the largest components are non-contact (i.e., the failure to ever reach a person at the location or phone number designated as part of the sample) and refusal (the result of active or passive activities to avoid completing the survey).

Measuring non-response bias is a difficult task, simply by definition. In its most elemental form, the question is: How do the people who did not complete the survey differ from those who did? Thus, by definition, less is known about those who did not complete the survey because... they did not complete the survey.

There have been a variety of excellent summaries of the research on non-response and potential bias from non-response in surveys, the latest of which is *Public Opinion Quarterly, Special Issue: Non-Response Bias in Household Surveys*.³ The introduction to that volume summarizes far better than will be attempted here the past and current state of research on non-response.

The *POQ* special issue includes a report on the most recent major experiment for gauging the impact of extra effort to complete

² OMB Clearance Number, 3137-0070, expiration date 06.30.2010.

³ *Public Opinion Quarterly, Special Issue: Non-Response Bias in Household Surveys*, 2006, Vol. 70, No. 5.

telephone interviews.⁴ This updated an earlier effort to measure the value to extraordinary attempts to reach nonrespondents.⁵ A central feature of each of these two experiments was to compare a standard RDD survey completed over a five-day period, to a RDD survey using the same questionnaire, the same sample design, the sample field house but a calling period that stretched for more than six months. PSRAI participated in each of these experiments.

It is well accepted that repeated callbacks are effective in reducing non-contact non-response (cites omitted) and that large interviewer workloads can reduce the ability interviewers to make such callbacks (cites omitted).⁶

The Design

Based on those models, this experiment was designed with three elements.

First, in line with Keeter, et. al., an analysis was conducted of survey results using the original 10-call design by examining the amount of effort actually required to get results. Respondents who broke off the original interview, households to which phone numbers were dialed seven or more times or where respondents originally refused to be surveyed were identified for this analysis. For the cases in the dataset, an effort variable was computed that represents the amount of effort it took to complete the interview, including call attempts, refusal conversions, etc. Then survey results were analyzed by the effort variable.

Second, the sample telephone numbers, based on their area code and exchange, can be matched to county with a good degree of reliability. PSRAI has a database with demographic information for all counties in the United States, including percent minority households, average household income and population density. These community-level demographics were appended to all cases in the sample and then the completed interviews were compared to the refusals on these measures. This provides an indication of what kinds of communities are under- and over-represented in our sample. And the next step was to compare substantive responses of the over- and under-represented areas to get a measure of non-response bias.

⁴ *Gauging the impact of Growing Nonresponse on Estimates from a National RDD Telephone Survey*, Keeter, et. al. *POQ*, 70: 759-779.

⁵ *Consequences of Reducing Nonresponse in a Large National Telephone Survey*, Keeter, et. al., *POQ* 64:125-48.

⁶ *Nonresponse Rate and Nonresponse Bias*, by Robert M. Groves, *POQ*, Vol. 70: 666.

The third and largest element of the non-response bias analysis called for drawing a sample of working telephone numbers from among those which have reached the maximum of 10 calls and then calling each number back another 10 times or until a terminal disposition is reached. The goal was that approximately 1500 telephone numbers would be pulled for this intensive effort.

While the comparison is not perfect, this is roughly the number of telephone numbers dialed beyond 10 times in the Pew Research Center infinite call design project reported on by Keeter.⁷ That effort obtained approximately 170 new interviews. Since that study was conducted several years ago, it was expected that this effort would obtain somewhat fewer completes.

It was not known if this design would provide statistically significant information. The much larger Keeter et.al. study (where the most rigorous portion of the study involved three times the resources of the less rigorous) produced only a handful of statistically significant differences. The level of effort for this current project was chosen to provide assistance in looking at the main survey results, without draining resources from the principal effort.

The Main 2007 Survey

The survey is based on data from telephone interviews conducted by Princeton Survey Research Associates International between June 27 to September 4, 2007, among a sample of 2,796 adults, 18 and older. For results based on the total sample, one can say with 95% confidence that the error attributable to sampling is plus or minus 2.5 percentage points.

Two separate samples were used for interviewing. The main sample yielded 2,063 interviews. This RDD consolidated sample was drawn disproportionately from areas in the country with higher than average numbers of African-American and Latino residents. In order to increase the number of low-access Internet users in our sample, an additional 733 interviews were conducted with respondents from low-access households that were identified in previous PIAL surveys.

⁷ *Gauging the impact of Growing Nonresponse on Estimates from a National RDD Telephone Survey*, Keeter, et. al. *POQ*, 70: 759-779.

Sample was released for interviewing in replicates, which are representative subsamples of the larger sample. Using replicates to control the release of sample ensures that complete call procedures are followed for the entire sample. At least 10 attempts were made to complete an interview at sampled households. Calls were staggered over times of day and days of the week to maximize the chance of making contact with potential respondents. Each household received at least one daytime call in an attempt to find someone at home. In each contacted household, interviewers asked to speak with the youngest male currently at home. If no male was available, interviewers asked to speak with the youngest female at home. This systematic respondent selection technique has been shown to produce samples that closely mirror the population in terms of age and gender.

Non-response in telephone interviews produces some known biases in survey-derived estimates because participation tends to vary for different subgroups of the population, and these subgroups are likely to vary also on questions of substantive interest. In order to compensate for these known biases, the sample data are weighted in analysis.

Response Rate for The Main Survey

PSRAI calculates a response rate for the total survey as the product of three individual rates: the contact rate, the cooperation rate, and the completion rate. Of the residential numbers in the sample, 83 percent were contacted by an interviewer and 36 percent agreed to participate in the survey. Seventy-three percent were found eligible for the interview. Furthermore, 90 percent of eligible respondents completed the interview. Therefore, the final response rate is 27 percent.

Following is the full disposition of all sampled telephone numbers:

Table 1: Combined Sample Disposition	
33821	Total Numbers Dialed
12214	Business / Government
1687	Computer/Fax
13	Cell phone
4071	Other not working
1718	Additional projected not working
14118	Working numbers
41.7%	Working Rate
496	No Answer
77	Busy
1585	Answering Machine
203	Other Non-Contacts
11757	Contacted numbers
83.3%	Contact Rate
814	Callbacks
6675	Refusal before eligibility status known
4268	Cooperating numbers
36.3%	Cooperation Rate
899	Language Barrier
247	Screen-outs for callback sample
3122	Eligible numbers
73.1%	Eligibility Rate
326	Refusal - Refusal after case determined eligible
2796	Completes
89.6%	Completion Rate
27.1%	Response Rate

Section I: The Impact of Interviewing Effort

The following analysis was done to assess the distribution of completed interviews by various demographic indicators in relation to the amount of effort expended to complete them. Respondents who broke off the original interview, households to which phone numbers were dialed seven or more times or where respondents originally refused to be surveyed were identified for this analysis.

Definition of Effort Variable

A three-category variable was computed to aid in the analysis of the amount of effort it took to complete interviews.

- Phone numbers that had been dialed six or more times in the original sample **and** where potential respondents refused to be surveyed, were defined as “hard,” meaning the actual interviewing effort was highest to complete an interview.
- Phone numbers that had been called five or fewer times in the original sample and where no potential respondents refused to participate, were defined as “easy,” meaning the actual interviewing effort was lowest to complete an interview.
- All other phone numbers were defined as “medium” effort, including respondents that were called six or more times **or** had refused to participate in an interview, but not both.

The hypothesis is that, among various demographic categories, it is less difficult to convert incomplete or as-yet-started interviews among respondents classified as easy than it is to complete interviews with respondents in the medium or hard categories.

Impacts of Effort on Personal-level Demographic Distributions

The apparent effect of expending more effort to complete interviews was greatest in certain demographic segments of respondents. Table 2 compares sample demographics of respondents according to the amount of effort it took to complete an interview. The population parameters are also represented in the table, as a point of reference.

The gender distribution among the effort categories is virtually uniform, reflecting the corresponding distribution of gender in the original sample. Expending greater efforts proved positive in reaching younger respondents, particularly those 18-29 years of age. Older respondents (65 years of age or older) who were defined as easy to reach were overrepresented as compared to the population parameter; completion rates are significantly, and precipitously lower for medium- or hard-to-reach respondents.

The greater effort expended to complete interviews with whites was most useful in re-contacting respondents in the easy group (who hadn't refused participation or were called fewer than six times previously) where 69% were white.

Those respondents in the medium or hard categories were less likely to be white. Conversely, more effort led to a greater number of completed interviews with African Americans. There was little difference by effort in reaching Hispanics.

Table 2: Individual-level Demographics By Effort Level

	Easy	Medium	Hard	Population Parameter ₂
	(n=1327)	(n=508)	(n=228)	
<u>Gender</u>				
Male	39%	42%	38%	48%
Female	<u>61%</u>	<u>58%</u>	<u>62%</u>	<u>52%</u>
	100%	100%	100%	100%
<u>Age</u>				
18-29	10%	15%	17%	21%
30-49	32%	32%	38%	39%
50-64	28%	29%	25%	24%
65+	28%	22%	17%	16%
Refused	<u>3%</u>	<u>3%</u>	<u>4%</u>	<u>NA</u>
	100%	101%	101%	100%
<u>Education</u>				
HS grad or less	40%	42%	41%	50%
Some college	26%	25%	24%	24%
College Grad.	35%	32%	33%	26%
Refused	<u>*1</u>	<u>1%</u>	<u>2%</u>	<u>NA</u>
	101%	100%	100%	100%
<u>Race/Ethnicity</u>				
White, not Hispanic	69%	62%	59%	71%
Black, not Hispanic	16%	21%	25%	11%
Hispanic	9%	11%	10%	12%
Other, not Hispanic	5%	4%	6%	6%
Refused	<u>2%</u>	<u>3%</u>	<u>1%</u>	<u>NA</u>
	101%	101%	101%	100%
<u>Employment Status</u>				

Employed FT or PT	53%	59%	59%	64%
Retired	30%	24%	20%	16%
Not employed	13%	12%	17%	14%
Disabled/Student/ Other	3%	3%	5%	5%
Refused	<u>*1</u>	<u>1%</u>	<u>0%</u>	<u>NA</u>
	99%	99%	101%	100%

¹ An asterisks indicated less than one percent.

² The population parameters were derived from the Census Bureau's 2006 Annual Social and Economic Supplement survey. The analysis included only continental U.S. households with a telephone.

It proved more difficult to complete interviews with respondents in the medium and hard categories who were employed full- or part-time, compared to respondents who said they were retired. Where the greater effort paid off in higher completion rates among the employed in the medium and hard categories, extra efforts were less fruitful in completing interviews with retired persons in the medium and hard categories (30% among easy vs. an aggregate of 22% for medium and hard).

Impacts of Effort on Household-level Demographic Distributions

There were no significant differences among the effort categories based on household characteristics, such as income and the presence of children in a household.

The level of interviewing effort was not related to reaching an appropriate proportion of respondents who make more than \$100,000 a year. Among households with children age 18 or less, effort did make a difference. About half of the medium (50%) and hard-to-reach respondents (51%) came from households with one or more children, compared with 45% for the lower levels of effort.

Table 3: Household Demographics by Effort Level

	Easy	Medium	Hard	Population Parameter ₂
	(n=1327)	(n=508)	(n=228)	
<i>Annual household income</i>				
Less than \$40K	35%	34%	36%	34%
\$40k - less than \$100K	29%	34%	29%	43%
\$100K or more	13%	13%	13%	23%
Refused	<u>23%</u>	<u>20%</u>	<u>22%</u>	<u>NA</u>
	100%	101%	100%	100%
<i>Children in household</i>				
None	54%	49%	50%	61%
One or more	45%	50%	51%	38%
Refused	<u>1%</u>	<u>1%</u>	<u>0%</u>	<u>NA</u>

100% 100% 100% 100%

¹ An asterisks indicated less than one percent.

Impacts of Effort on Community-level Demographic Distributions

We also tested to see if more effort helped complete interviews in harder-to-reach communities. Typically in RDD telephone samples, households in heavily populated urban areas are under-represented. These households are also more likely to have minority residents. We appended two variables onto the dataset which tell us about the communities where the respondents reside. Both variables came from matching area code/ exchange combinations to 2000 Census tract-level data.

The first variable we appended was population density. This is a 5-category variable that divides the population into five equal-sized groups. The highest-density group represents the most densely populated areas that contain 20 percent of the continental U.S. population. The lowest-density group represents the least densely-populated areas where 20 percent of the population resides.

The second variable we appended measures the percent of the community's population that is minority. We define minority as either African-American or Latino.

Table 4 compares the three sub-samples in relation to population density and the percentage of a community's population that is minority. There were no significant differences among completion rates with respect to the levels of effort expended to get respondents to participate. There was no significant variation in reaching respondents in high-density areas with respect to the level of effort expended to complete interviews.

There were no significant differences among the effort categories they relate to the proportion of minorities in the sub-samples.

Table 4: Community Demographics by Effort Level

	Easy (n=1327)	Medium (n=508)	Hard (n=228)	Population Parameter
<i>Population Density</i>				
1 - Lowest	23%	24%	19%	20%
2	19%	16%	22%	20%
3	18%	18%	16%	20%
4	19%	19%	21%	20%
5 - Highest	<u>22%</u>	<u>24%</u>	<u>22%</u>	<u>20%</u>
	100%	101%	100%	100%
Percent Minority^{1,2}	30%	33%	31%	23%

¹ Minority is defined as either African-American or Latino.

² Both samples contained higher than average minority densities because the RDD sample was designed to over-represent both African-Americans and Latinos.

Impacts of Effort on Substantive Question Response Distributions

Of particular interest to this, or any, research is the possible effect that extra effort might have on the substantive results. For this analysis, we have selected a subset of questions to investigate the effects of extra effort on survey results.

The first series of questions we analyzed was the Q2 series, about visiting various local institutions in the past 12 months. There is little variation among the affirmative responses in the Q2 series. None of the differences are statistically significant.

Table 5: Visiting Local Institutions By Effort

	Easy (n=1327)	Medium (n=508)	Hard (n=228)
<i>During the past 12 months, have you gone to a...</i>			
Local public library (Q2a)	51%	55%	52%
Local place where you can use a computer for free (Q2b)	27%	32%	27%
Government office or agency in person (Q2c)	43%	42%	42%
Local courthouse (Q2d)	34%	36%	31%

Comparing the results of a series of questions about computer and Internet use among respondents in the sub-groups, we observed no

statistically significant differences in responses among effort segments. Table 6 compares results for this series, which includes Q5 to Q8.

Table 6: Computer and Internet Use By Effort			
	Easy	Medium	Hard
	<i>(n=1327)</i>	<i>(n=508)</i>	<i>(n=228)</i>
<i>Percent of adults who...</i>			
Use a computer on at least an occasional basis (Q5)	74%	77%	75%
Use the Internet at least occasionally (Q6a)	71%	74%	75%
Send and receive email at least occasionally (Q6b)	67%	69%	68%
Are Internet Users₁	72%	75%	76%
<i>Percent of Internet users who...</i>			
Go online from home (Q7a)	94%	92%	93%
Go online from work (Q7b)	60%	64%	59%
Have low-access to the Internet₂	18%	22%	21%

¹ Internet users are defined as people who report either using the Internet or sending/receiving email at least occasionally.

² Low access Internet users are defined as those who have dial-up access from home.

For this analysis, we then examined differences in the Q9 series, which asked about ten situations or decisions that people might have faced in the past two years. These questions are one of the two core groups of substantive questions in the survey. Table 7 summarizes the responses to this series.

There were statistically significant differences between the three subgroups on just two of the ten items on the list. The toughest-to-reach respondents were more likely than others in the easy and medium categories to have made a decision about schooling or education (46% vs. 40% for medium and 34% for easy). Age and parental status are each likely the factors affecting these results. Younger respondents are typically harder to reach, appear in the extra-effort sample in higher proportions than in the original study, and are inherently more likely to have recently made decisions about their education.

The greater likelihood among easy (48%) and medium (51%) category respondents to have dealt with a serious illness or health condition is attributable to older and retired respondents being easier to contact.

Table 7: Decisions and Situations By Effort			
	Easy	Medium	Hard
	<i>(n=1327)</i>	<i>(n=508)</i>	<i>(n=228)</i>
<i>Percent of people who in the last two years have...</i>			
Made a decision about schooling or education (Q9a)	34%	40%	46%
Dealt with a serious illness or other health condition (Q9b)	48%	51%	40%
Needed information about Medicare, Medicaid or food stamps (Q9c)	26%	25%	25%
Changed jobs, retired, or started a business (Q9d)	23%	24%	22%
Been involved in a criminal matter, a lawsuit, or other legal action (Q9e)	9%	10%	13%
Needed information about Social Security or military benefits (Q9f)	24%	22%	18%
Needed information about property or income taxes (Q9g)	35%	36%	32%
Become a citizen or helped someone through the immigration process (Q9h)	4%	5%	5%
Looked for help from local government (Q9i)	14%	16%	15%

Table 7: Decisions and Situations By Effort			
	Easy	Medium	Hard
Wanted information about voter registration (Q9j)	20%	20%	16%

The next step was to analyze the results from the Q11 series, which asks about the sources people use to get information or assistance. The Q11 series was the second major substantive series of questions in the survey. Here the results were generally uniform among the effort categories. There were slight statistically significant differences in terms of those who used the Internet to find information or assistance in solving their problems. Medium- and hard-to-reach respondents were slightly more likely than the easy-to-reach respondents to use the Internet (57% for hard/medium vs. 53% for easy). This correlates with the greater usage of the Internet by younger people, who are, in turn, harder to reach.

Table 8: Sources for Information and Assistance by Effort			
	Easy	Medium	Hard
	<i>(n=1327)</i>	<i>(n=508)</i>	<i>(n=228)</i>
<i>Percent of people who used the following sources...</i>			
Friends and family (Q11a)	41%	43%	45%
Professionals (e.g., doctors, lawyers) (Q11b)	54%	57%	47%
The Internet (Q11c)	53%	58%	56%
Newspapers, magazines or books (Q11d)	37%	39%	35%
Television or radio (Q11e)	18%	19%	19%
Public library (Q11f)	12%	16%	13%
Local place with free computer access (Q11g)	9%	12%	9%
Government office or agency (Q11h)	34%	35%	29%
Another source not previously mentioned (Q11i)	17%	16%	18%

Section II: Community Demographics and Refusals

The second part of the analysis of potential non-response bias includes an examination of final sample dispositions of numbers dialed by community to assess over- or under-representation in the original sample. The disposition categories include completed interviews (completes), a combination of refusals and callbacks (refusals/callbacks), and numbers dialed where no potential respondent was contacted (non-contacts).

Community-level demographics have been appended to all cases in the sample by matching census tract-level data to telephone exchanges. Appended demographics analyzed are percent minority households, average household income, community types and region of the United States.

Variation by Region and Community Type

There are multiple dimensions of variation revealed when assessing non-response based on sample demographics describing location and types of communities where respondents live. Table 9 compares the effect of region, community type and whether those who were sampled live in a Metropolitan Statistical Area (MSA)⁸ or not, on response.

	Completes (n=2063)	Refusals/ Callbacks (n=6213)	Non- Contacts (n=2452)	Total
<i>Region</i>				
Northeast	16%	56%	28%	100%
Midwest	23%	53%	24%	100%
South	18%	53%	29%	100%
West	15%	51%	34%	100%
<i>Community Type</i>				
Urban	15%	52%	33%	100%
Suburban	18%	53%	28%	99%
Rural	25%	56%	19%	100%
<i>Metros</i>				
Metropolitan Statistical Areas	17%	53%	31%	101%
Non-MSA	25%	56%	19%	100%

⁸ MSAs are delineated on the basis of a central urbanized area—a contiguous area of relatively high population density. The counties containing the core urbanized area are known as the central counties of the MSA. Additional surrounding counties (known as outlying counties) can be included in the MSA if these counties have strong social and economic ties to the central counties as measured by commuting and employment. Note that some areas within these outlying counties may actually be rural in nature.

As can be seen in the above table, differences in response are mainly seen among geographical regions represented in the sample.

A higher relative proportion of interviews with respondents from the Midwest were completed (23%), as compared to just 15 percent in the West. Fully one third of the sample in the West region was unable to be contacted by PSRAI interviewers. Similarly, interviewers were unable to contact a potential respondent for 33% of the sample that came from urban areas. The West sample itself was significantly skewed toward urban communities, comprising 54% of the total sample for the region.

The combined proportion of refusals and callbacks was relatively uniform among the regions. A higher proportion of completed interviews can be observed among those sampled in rural areas and outside of Metropolitan Statistical Areas.

Income, Minority Density and Response

Table 10 compares the disposition categories with average household income and percentage density of Hispanics and African-Americans within sample blocks used for interviewing.

There is very little difference among income distributions for the three sample segments. Between 35% and 40% of each segment came from areas with lower household incomes and less than 10 percent of each group came from the highest income areas.

Table 10: Community Demographics By Status			
	Completes	Refusals/ Callbacks	Non- Contacts
	<i>(n=2063)</i>	<i>(n=6213)</i>	<i>(n=2452)</i>
<u>Average Household Income</u>			
Less than \$40K	35%	38%	40%
\$40k - less than \$100K	59%	57%	54%
\$100K or more	<u>6%</u>	<u>5%</u>	<u>7%</u>
Total	100%	100%	101%
<u>Projected Hispanic incidence</u>			
0-8%	63%	53%	46%
9-19%	11%	10%	12%
20-34%	15%	20%	21%
35%+	<u>12%</u>	<u>17%</u>	<u>21%</u>
Total	101%	100%	100%
<u>Projected African-American Incidence</u>			
0-9%	59%	56%	56%
10-29%	18%	19%	20%
30-49%	10%	11%	12%
50%+	<u>13%</u>	<u>14%</u>	<u>13%</u>
Total	100%	100%	101%

Interviews were easier to conduct in areas with fewer Hispanic households. Nearly two-third of all completed interviews (63%) came from areas with the lowest incidence of Hispanic households. Only about half of the sample in the other groups came from these low-incidence Hispanic areas (53% of refusals/callbacks and 46% of non-contacts). The same trend is not seen when comparing the high and low-density African-American areas. Completes, refusals and non-contacts were distributed about the same across the African-American strata.

Section III: Extra Effort, Extra Calls

In an effort to analyze potential non-response bias in survey results, a sample of 1,500 phone numbers that had been called 10 times and were still live numbers were put into a new project and dialed up to ten more times to try and complete an interview. The numbers that were included in the extra-effort dialing fell into three categories.

- **Non-contacts:** Numbers that yielded no contact with a person. These numbers were some combination of no answer/ busy/ answering machines for all attempts.
- **Refusals:** Numbers that yielded a refusal on one of the first ten attempts but were not converted to a completed interview.⁹
- **Break-offs:** Numbers where respondents started an interview but did not finish.

An additional 13,742 calls were made to the extra-effort sample and an additional 84 interviews were completed. This translates into one completed interview for every 164 call attempts. During the first ten calls, one interview was completed for every 78 calls made.¹⁰

Effect of Extra Effort on Sample Disposition Outcomes

It is difficult to complete an interview once a phone number has been dialed 10 times without success. Table 11 compares the outcome of the extra-effort numbers at the 10th attempt (the columns) with the outcome after 20 or more calls (the rows).

	Status After First 10 Attempts...				
	<u>Non-contact</u>	<u>Refusal</u>	<u>Break-off</u>	<u>Callback</u>	<u>Total</u>
<i>After 10 Extra calls...</i>					
Non-contact	676	75	13	37	801
Refusal	143	201	18	69	431
Break-off	3	3	1	2	9
Callback	19	24	3	16	62
Complete	33	25	7	19	84
<u>Other</u>	<u>75</u>	<u>28</u>	<u>3</u>	<u>7</u>	<u>113</u>
Total	949	356	45	150	1500
Percent converted	3%	7%	16%	13%	6%

As can be seen from the table, the success rate was low, with only 6 percent of the extra-effort cases yielding completed interviews. The conversion rate was higher for the break-offs and the callbacks, with

⁹ Hard refusals were excluded from the extra effort study. These are phone numbers where the potential respondents have refused to cooperate in no uncertain terms.

¹⁰ The difference in calling productivity between the two samples is even more pronounced if you consider that many of the non-working phone numbers were identified on the first attempt before the extra effort even started.

16% and 14% converted respectively. The numbers that were non-contacts for the first ten attempts had the lowest conversion rate (3%). Seven percent of the refusals were converted after the tenth attempt.

Effect of Extra Effort on Person-level Demographics¹¹

One way to gauge potential non-response bias is to compare the demographics of the respondents interviewed with the extra effort to those interviewed with no extra effort. Table 12 compares basic sample demographics of the two groups. The population parameters are also represented in the table to put the numbers in context.

Table 12: Personal Demographics by Extra Effort			
	No	Extra	Population
	Extra	Extra	Parameter ₂
	Effort	Effort	
	(n=2063)	(n=84)	
<u>Gender</u>			
Male	40%	39%	48%
Female	<u>61%</u>	<u>61%</u>	<u>52%</u>
	101%	100%	100%
<u>Age</u>			
18-29	12%	17%	21%
30-49	32%	38%	39%
50-64	28%	27%	24%
65+	25%	13%	16%
Refused	<u>3%</u>	<u>5%</u>	<u>NA</u>
	100%	100%	100%
<u>Education</u>			
HS grad or less	40%	30%	50%
Some college	25%	23%	24%
College Grad.	34%	44%	26%
Refused	<u>1%</u>	<u>4%</u>	<u>NA</u>
	100%	100%	100%
<u>Race/Ethnicity</u>			
White, not Hispanic	66%	58%	71%
Black, not Hispanic	18%	17%	11%
Hispanic	10%	14%	12%
Other, not Hispanic	5%	7%	6%
Refused	<u>2%</u>	<u>4%</u>	<u>NA</u>
	101%	100%	100%
<u>Employment Status</u>			
Employed FT or PT	55%	66%	64%
Retired	28%	17%	16%
Not employed	14%	13%	14%
Disabled/Student/			
Other	4%	2%	5%
Refused	<u>*₁</u>	<u>2%</u>	NA
	101%	100%	100%

¹¹ Unweighted data was used for all comparisons in this report.

¹ An asterisks indicated less than one percent.

² The population parameters were derived from the Census Bureau's 2006 Annual Social and Economic Supplement survey. The analysis included only continental U.S. households with a telephone.

The gender distribution of the two samples was almost identical, 39 percent and 40 percent males. As is common in most telephone surveys, males are under-represented. There was a difference in the age distribution of the two samples with the extra-effort sample doing a slightly better job of reaching younger respondents, especially the 30-49 year old group. Additionally, the extra effort yielded a more appropriate proportion of older respondents (13%) than the regular sample (25%).

Both samples over-represented college graduates and under-represented people with less education. However, on the whole the regular sample got a slightly better overall education distribution.

The race/ethnicity of the two samples was comparable. However, the extra effort did pay off with reaching more Hispanic respondents than the standard sample. The extra-effort sample also reached more employed respondents than the standard sample (66% vs. 55%) and did not over-represent retired people as did the standard sample.

Effect of Extra Effort on Household-level Demographics

It is also useful to look at household characteristics to get a sense of the kinds of households that might not be reached in a 10-call design, but get contacted beyond the tenth call. Table 13 compares the two samples on two household-level demographics – annual household income and the presence of children.

In terms of household income, the main difference is that the completes from the extra-effort sample yielded a higher level of refusals than the regular sample (32% vs. 22%). The presence of children in the household was about the same for the two samples, with 30 percent of the standard sample and 33 percent of the extra-effort sample reporting at least one child in the household.

Table 13: Household Demographics by Extra Effort			
	No Extra Effort	Extra Effort	Population Parameter²
	(n=2063)	(n=84)	
<u>Annual household income</u>			
Less than \$40K	35%	25%	34%
\$40k - less than \$100K	30%	32%	43%
\$100K or more	13%	11%	23%
Refused	<u>22%</u>	<u>32%</u>	<u>NA</u>
	100%	100%	100%
<u>Children in household</u>			
None	70%	66%	61%
One or more	30%	33%	38%
Refused	<u>*</u>	<u>1%</u>	<u>NA</u>
	100%	100%	100%

¹ An asterisks indicated less than one percent.

² The population parameters were derived from the Census Bureau's 2006 Annual Social and Economic Supplement survey. The analysis included only continental U.S. households with a telephone.

Effect of Extra Effort on Community-level Demographics

We also tested to see if extra effort helped get interviews in harder-to-reach communities. As explained earlier, we appended two variables onto the dataset which tell us about the communities where the respondents reside. The variables come from matching area code/exchange combinations to 2000 Census tract-level data.

The first variable was population density, a 5-category variable that divides the population into five equal-sized groups. The second variable measures the percent of the community's population that is minority. Again, we define minority as either African-American or Latino.

Table 14 compares the two samples with respect to these two community variables. The extra effort did get more respondents from high-density areas. For example, 55 percent of the extra-effort completes came from the two highest density strata compared with 41 percent of the regular sample. The extra effort also yielded a higher proportion of minority interviews as well (34% vs. 30%).

Even though the standard sample distributions of these two variables are closer to the population parameters, this is only the result of the disproportionate sample design that was employed for this survey. In surveys with standard RDD sample designs, minority households along with households in densely-populated urban areas are usually under-represented.

	No Extra Effort (n=2063)	Extra Effort (n=84)	Population Parameter
<i>Population Density</i>			
1 - Lowest	23%	16%	20%
2	19%	20%	20%
3	17%	10%	20%
4	19%	26%	20%
5 - Highest	<u>22%</u>	<u>29%</u>	<u>20%</u>
	100%	100%	100%
Percent Minority_{1,2}	31%	34%	23%

¹ Minority is defined as either African-American or Latino.

² Both samples contained higher than average minority densities because the RDD sample was designed to over-represent both African-Americans and Latinos.

Effect of Extra Effort on Selected Substantive Questions

Of particular interest to this, or any, research is the possible effect that extra effort might have on substantive results. For this analysis, we have selected a subset of questions to investigate the effects of extra effort on survey results.

The first series of questions we analyzed was the Q2 series which asked respondents if they have gone to various local institutions in the past 12 months. As can be seen in Table 15, the level of “Yes” responses to all four items in the two samples is virtually identical.

	No Extra Effort (n=2063)	Extra Effort (n=84)
<i>During the past 12 months, have you gone to a...</i>		
Local public library (Q2a)	52%	51%
Local place where you can use a computer for free (Q2b)	28%	29%
Government office or agency in person (Q2c)	43%	42%
Local courthouse (Q2d)	34%	37%

We also tested a series of questions that asked about computer and Internet use (Q5 to Q8). Table 16 compares results for these questions. There are no significant differences between the two samples' responses with the exception of Q7b. Those reached during the extra-effort portion of the study were more likely than regular sample respondents to report going online from work (78% vs. 61%). This difference is likely due to the fact that employed people are more difficult to reach and appear more often in the extra-effort sample.

Table 16: Computer and Internet Use		
	No Extra Effort	Extra Effort
	<i>(n=2063)</i>	<i>(n=84)</i>
<i>Percent of people who...</i>		
Use a computer on at least an occasional basis (Q5)	75%	79%
Use the Internet at least occasionally (Q6a)	72%	77%
Send and receive email at least occasionally (Q6b)	68%	68%
Are Internet users¹	73%	76%
<i>Percent of Internet users who...</i>		
Go online from home (Q7a)	93%	89%
Go online from work (Q7b)	61%	78%
Have low-access to the Internet²	19%	14%

¹ Internet users are defined as people who report either using the Internet or sending/receiving email at least occasionally.

² Low access Internet users are defined as those who have dial-up access from home.

We also investigated differences in the Q9 series which asked about ten situations or decisions that people might have faced in the past two years. Table 17 summarizes the responses to this series.

There were statistically significant differences between the two samples on four of the ten items in the list. Respondents interviewed during the extra-effort portion of the study were less likely than regular sample respondents to have dealt with a serious illness (33% vs. 48%). They are also less likely to report having needing information about Medicare, Medicaid or food stamps (17% vs. 25%). These two differences are probably due to the age of respondents. Younger respondents are typically harder to reach and therefore appear in higher proportions in the extra-effort sample.

Additionally, the extra-effort respondents are more likely to report either becoming a citizen or helping someone else through the immigration process (11% vs. 4%), and they are less likely to report having looked for help from their local government (7% vs. 15%).

Table 17: Decisions and Situations		
	No Extra Effort	Extra Effort
	<i>(n=2063)</i>	<i>(n=84)</i>
<i>Percent of people who in the last two years have...</i>		
Made a decision about schooling or education (Q9a)	37%	41%
Dealt with a serious illness or other health condition (Q9b)	48%	33%
Needed information about Medicare, Medicaid or food stamps (Q9c)	25%	17%
Changed jobs, retired, or started a business (Q9d)	23%	19%
Been involved in a criminal matter, a lawsuit, or other legal action (Q9e)	9%	7%
Needed information about Social Security or military benefits (Q9f)	23%	20%
Needed information about property or income taxes (Q9g)	35%	37%
Become a citizen or helped someone through the immigration process (Q9h)	4%	11%
Looked for help from local government (Q9i)	15%	7%
Wanted information about voter registration (Q9j)	20%	14%

Finally, we looked at the Q11 series to see if there were any differences between samples in the sources people use in getting information or assistance. Results for the two samples were very similar with the exception of Internet use to find information or assistance in solving problems. More people in the extra-effort sample reported using the Internet to get information about or assistance with their recent decision or situation (68% vs. 55%).

Table 18: Sources for Information and Assistance		
	No Extra Effort	Extra Effort
	<i>(n=2063)</i>	<i>(n=84)</i>
<i>Percent of people who used the following sources...</i>		
Friends and family (Q11a)	42%	34%
Professionals (e.g., doctors, lawyers) (Q11b)	54%	51%
The Internet (Q11c)	55%	68%
Newspapers, magazines or books (Q11d)	37%	36%
Television or radio (Q11e)	18%	14%
Public library (Q11f)	13%	15%
Local place with free computer access (Q11g)	10%	14%
Government office or agency (Q11h)	34%	34%
Another source not previously mentioned (Q11i)	17%	22%

A Final Note about Potential Bias

In conclusion, survey results from the extra-effort study varied only occasionally from results in the base study.

The extra interviewing effort had the expected impact: that is, those who are harder to reach in general were those reached with extra effort. Younger adults, working adults and those with college degrees are harder to reach and thus were a larger share of the extra-effort completes. In terms of substantive questions, the impact of the extra effort parallels the demographic differences. Those reached with extra effort were more likely to say they have dealt with an education issue (reflecting the younger adults) and less likely to have dealt with a health issue or Medicare (reflecting the larger share of older adults in the normal sample).

Even when results were significantly different with the extra-effort group, the impact on overall results was small.

Consider that the sample of 1,500 extra-effort numbers only yielded 84 completed interviews. There were only about 3,000 numbers that qualified for the extra-effort study, so even if all of them had been dialed an additional 10 times that would have added a total of approximately 170 completes to our main sample of over 2,000. Even if the extra-effort results were different than the main sample, there simply would not be enough of them to move overall survey results.