June 24, 2002

American Community Survey Research and Evaluation Program Discussion Paper

Evaluating the Feasibility of Conducting Mailout and Mailback Operations in Puerto Rico in 2003

Theresa Leslie, Michael Starsinic, and Anthony Tersine, DSMD

Abstract: The American Community Survey (ACS) uses mail, computer assisted telephone interviewing (CATI), and computer assisted personal visit interviewing (CAPI) to collect the survey data. In November and December of 2001, ACS staff conducted a test to assess the feasibility of using "mail" as a data collection methodology in Puerto Rico. This is the first time the Census Bureau attempted to collect data using mailout methods in Puerto Rico. The results from this test show that we received mail responses from 27.8 percent of the addresses to which we mailed an ACS mail questionnaire. This report also provides results on the deliverability rate and the quality of the deliverability indicator on the contractor's file.

Keywords: Puerto Rico, addresses, mail

Document Number: D-33

---

1. Introduction

The American Community Survey (ACS) is proposed as a replacement for the Census long form. Full implementation of the ACS is scheduled for 2003. In preparation for full implementation, research and development activities began in 1996. The Continuous Measurement program was started in four test sites and expanded to 36 counties in 1999. In addition, the Census 2000 Supplementary Survey (C2SS) was conducted as part of Census 2000 in another 1,203 counties. In 2001, we conducted the 2001 Supplementary Survey. All the tests use ACS methods and questionnaires. The ACS will use mail, computer assisted telephone interviewing (CATI), and computer assisted personal visit interviewing (CAPI) to collect the survey data.

We conducted the Puerto Rico Test concurrent with the 2001 Supplementary Survey. In November and December of 2001 staff conducted a test to assess the feasibility of using "mail" as a data collection methodology in Puerto Rico. This is the first time the Census Bureau attempted to mail to addresses in Puerto Rico to conduct data collection. The update/leave/mailback data collection methodology was used for Census 2000. Census enumerators delivered questionnaires to each housing unit and the householders mailed the completed questionnaires back to the Census Bureau. While at the housing unit, the enumerators updated the mailing address on listing pages and verified or spotted the location of the unit on census maps, updating the Puerto Rico Master Address File (MAF). The base MAF for Puerto Rico was created as the result of the Address Listing operation which took place island wide in 1998. The ACS would like to use the three modes of data collection used stateside in Puerto Rico. This test therefore was designed to assess if mailout/mailback methods would work.

This study answers the following key questions:

A. What proportion of the sample addresses were determined to be unmailable (by our edit) and undeliverable (by the U.S. Postal Service or USPS)?

B. What were the mail response rates? How did they vary by stratum?

C. How effective is the contractor's deliverability flag (DLV) at identifying addresses on the Puerto Rico MAF that could not be delivered by the USPS?

D. How cooperative were respondents who received a questionnaire?

E. What do these data tell us about cost modeling assumptions for Puerto Rico for 2003?

2. Background

A. Puerto Rico Master Address File (MAF)

The addresses for the Puerto Rico MAF were listed during a Puerto Rico address listing operation in 1998 and later updated by census enumerators when delivering Census 2000 questionnaires. Stateside address listing pages and procedures were modified to list the addresses in Puerto Rico because Puerto Rico address conventions differ from stateside conventions. For example, urbanization[1] codes could appear in the address field, thus, the address could require three instead of two lines of information which is standard for stateside addresses. For example, a typical address in Puerto Rico is:

Name
Urbanization name
House Number and street name
City, State, and ZIP+4

The Decennial Systems and Contracts Management Office (DSCMO) had problems processing the keyed listing pages from the Address Listing operation in Puerto Rico. The keyed files had a 60 character address field that could contain a city-style address or a location description. The stateside files also had a flag, “A/D”, set by the lister that indicated which it was. In the U.S., census enumerators set the flag to “A” for a city-style address or “D” for a location description. In Puerto Rico, the flag was "D/L", and census enumerators set the flag to "D" for city-style address and "L" for location description. When the DSCMO processed the files for Puerto Rico, they initially assumed that the "D" in the flag identified a "location description", as it did in the U.S., but the "D" actually stood for address (the word for address in Spanish starts with a "D"). The DSCMO fixed this by re-processing the files.

Given the differences in address conventions for Puerto Rico, the stateside address standardizer could not be used to get the address information in the appropriate city-style address and location description fields on the MAF. The DSCMO and GEO decided to load the entire address field (city-style and location description information) in the location description field on the MAF. This processing decision continued for all address updating operations that the Census Bureau conducted in Puerto Rico after Address Listing. Because of this problem, there are no address records in the stateside-type layout for Puerto Rico on the MAF extracts used for this evaluation.

---

[1] Urbanization denotes an area, sector, or development within a geographic area. In addition to being a descriptive word, it precedes the name of the area. This descriptor, commonly used in Puerto Rican urban areas, is an important part of the addressing format of Puerto Rico, as it describes the location of a given street.

After Census 2000, a contractor was hired to create the MAF for Puerto Rico. The Geography Division contracted with SEEK Data to create the MAF for Puerto Rico because they have extensive data files for, and experience in, Puerto Rico. Using their own data, they improved the data in the house number, street name, and location description fields. They created an address label file for this test which included room for the name and three lines of address information. At the time of sample selection for this test, SEEK Data had only improved about 36 percent (510,854 of 1,327,250) of the addresses on the Puerto Rico MAF. Staff in the Demographic Statistical Methods Division (DSMD) compared the census mail response rates for these blocks to blocks where the addresses had not been improved. The DSMD also looked at how these addresses were distributed across Puerto Rico. They concluded that this 36 percent was not highly clustered in a few areas. Since the census mail response distributions were similar and the geographic coverage was adequate, the 510,854 addresses were determined to be acceptable as the sampling universe for this test.

The contractor's file includes a variable called deliverability code (DLV). The code values are:

1: "Exact match" to the Delivery Sequence File (DSF) address including ZIP+4.
2: "Equivocated match" to the DSF including ZIP+4. Some parts of the address do not match.
3: "Unmatched but deliverable"--The address doesn't match to the DSF but the full address is available.
4: "Undeliverable"--The address is missing critical address information.

This deliverability code was not used in determining mailable addresses. Part of this test is to determine if this variable is of value in determining postal deliverability.

B. The Sample

Staff in the DSMD selected a sample of 12,000 addresses from 510,854 addresses on the Puerto Rico MAF that SEEK Data had improved. Since the contractor was confident that the addresses parsed were representative of all addresses on the Puerto Rico MAF, only the 510,854 addresses parsed at the time of sample selection were in the sampling universe. Three strata were formed to represent (1) the San Juan Primary Metropolitan Statistical Area (PMSA), (2) other PMSAs or Metropolitan Statistical Areas (MSA) and (3) all other areas not in a PMSA or an MSA.

DSMD selected a stratified systematic sample proportional to the stratum size. The following table gives details of the stratum and sample sizes.

Table 1. Description of the Sample for the Puerto Rico Test

Stratum	Description	Addresses in Full MAF	Addresses in Universe for Sampling	Percentage of MAF	Sample size
1	In the San Juan PMSA	694,261	306,861	44.2%	6,277
2	In a MSA or a PMSA other than the San Juan PMSA	428,932	151,074	35.2%	3,878
3	Not in a MSA or a PMSA	204,057	52,919	25.9%	1,845
Total	Puerto Rico	1,327,250	510,854	38.5%	12,000

The sample was allocated evenly across two panels for November and December, 2001. Staff in the Demographic Surveys Division (DSD) edited the file of sample addresses to identify "mailable" addresses. "Mailable" addresses have address information in the first line of the address field (ADDLLINE1) and a complete Zip Code. P.O. Box addresses are considered "unmailable." The mailable address universe was used to label pre-notice letters, initial mailing packages and reminder cards. After check-in of mail returns, a second universe file (subset of the initial file) was defined to label the replacement mailing packages.

C. Adaptations to stateside materials were needed to mail out forms in Puerto Rico

In preparation for this test, staff in the Decennial Management Division translated a revised set of materials including the pre-notice letter, questionnaire mailing packages, and reminder card in Spanish, with specific additional revisions made for Puerto Rico. These materials were not available in Spanish for the ACS before this test. A specific job aid was developed for the Tucson Telephone Questionnaire Assistance Center's use to tell them procedurally how to handle calls from the Puerto Rico Test.

D. Implementing the Mail Data Collection Methodology

For this test, we implemented the complete stateside mailing strategy. The USPS attempted to deliver a pre-notice, initial questionnaire and reminder postcard during November and December of 2001 to all addresses in the 12,000 sample considered "mailable." If the USPS could not deliver the questionnaire, they returned it as "Undeliverable As Addressed (UAA)" to the National Processing Center (NPC) in Jeffersonville, IN for processing. We mailed a replacement questionnaire mailing package to mailable addresses that did not return the first questionnaire. We did not consider the USPS determination of undeliverable addresses when preparing for the second mailing; that is, we labeled a mailing package for the addresses the USPS identified as undeliverable when attempting to deliver the first questionnaire mailing package. The schedule below shows the dates that each mailing piece type left NPC for delivery by the USPS by panel for this test.

Table 2. Schedule for Mailing for Puerto Rico Test

Mailing Piece	Date Delivered by Panel	
	November	December
Advance Notice	October 25	November 21
Initial Questionnaire	October 29	November 26
Reminder Notice	November 1	November 29
Second Questionnaire	November 20	December 20

Addresses for which we did not receive a return by November 28 for the November panel and December 27 for the December panel were eligible for CATI. The results from the CATI follow-up operation are not included in this memorandum.

The NPC keyed the mail returns and DSD ran the automated edits on the keyed data to determine whether a case was eligible for telephone failed edit follow-up, but we did not conduct the follow-up.

3. Evaluation Methodology

A. Questions to Answer

1. Mailability and Deliverability
   What proportion of the sample addresses were determined to be unmailable (by our edit) and undeliverable (by the USPS in Puerto Rico)?
   How do these rates for Puerto Rico compare to rates for states in the 2001 Supplementary Survey?
2. Mail Response
   What were the mail response rates? How did they vary by sampling stratum?
   How do these rates for Puerto Rico compare to rates for states in the 2001 Supplementary Survey?
   How do the response rates for the sample panels compare?
3. Contractor's Deliverability Flag
   How effective is the contractor's deliverability flag (DLV) at identifying addresses on the Puerto Rico MAF that could not be delivered by the USPS?
   How many addresses identified as undeliverable by the contractor did our edit remove?
   Was the USPS able to deliver questionnaires to addresses considered mailable by our edit but undeliverable by the contractor (DLV=4)?
4. Cooperation
   How cooperative were respondents who received a questionnaire?
   How do these rates for Puerto Rico compare to rates for states in the 2001 Supplementary Survey?
5. Cost Modeling
   What do these data tell us about cost modeling assumptions for Puerto Rico in 2003?

B. Methodology

Most of the questions relating to the use of mailout/mailback methods will require the calculation of:

- Unmailable rate: Ratio of the number of addresses deemed "unmailable" by our edit to the total number of sample addresses,
- Mailable rate: Ratio of the number of addresses deemed "mailable" by our edit to the total number of sample addresses,
- Undeliverable rate: Ratio of the number of Undeliverable-as-Addressed (UAA) questionnaires to the mailout; a questionnaire was defined as a UAA if the USPS identified it as undeliverable during the first, second, or both mailings,
- Mail response rate: Ratio of the number of mail returned questionnaires to the mailout, and
- Cooperation rate: Ratio of the number of mail returned questionnaires to the number of addresses in the mailout that were not UAAs[2].

The DSMD produced unweighted rates for each of the sample months by strata and a combined, weighted estimate for the two months for the combined Puerto Rico test site. The total line in the tables in this report are the weighted estimates reflecting the 510,854 sampling universe.

C. Sources of Data

Control file information is the basis for much of the required evaluation data. Information on unmailable questionnaires comes from DSD. Questionnaires that are determined by the USPS to be undeliverable are returned to the NPC for check in. This check-in file is the source of data for the calculation of undeliverable rates. Information on deliverability (DLV) were provided by GEO. The list of variables used in the analysis and their source are included in Appendix 1.

D. Limitations

1. Our results assume that mail returns were completed by the correct sample address. We have no way of verifying this without conducting a reinterview operation. No such operation is planned at this time.

2. The questionnaires that the USPS could not deliver (Undeliverables) were destroyed accidently by staff in NPC. Therefore, we will not have the opportunity to learn anything about the reasons that the USPS was unable to deliver these mailing pieces.

4. Analysis

[2] Excluded only those UAAs without a corresponding mail return.

A. Mailability and Deliverability

What proportion of the sample addresses were defined as unmailable by our edit and undeliverable by the USPS in Puerto Rico?

All sample addresses were edited by DSD before they created the address label file used by the NPC staff to label the mailing pieces. The stateside edit was used for this test. The edit considered an address unmailable if no address information was available on the MAF or if the address was a post office box (even if additional address information was available). In this test, we hope to learn how to better define the edit for Puerto Rico addresses. Using the stateside edit for Puerto Rico addresses may be problematic given there are often banks of mailboxes at the entrances of "urbanizaciones." As the data in Table 3 show, almost eight percent of the sample addresses in the Puerto Rico test site were considered unmailable by the edit. The percent of unmailable addresses was higher for the "Other MSA" and the "Not MSA" stratum than the "San Juan PMSA."

The USPS identified undeliverable addresses when delivering the initial and replacement questionnaire mailing packages. If the USPS indicated that the mailing package was not deliverable when attempting to deliver either the initial or the replacement questionnaire mailing package, we considered the mailing piece undeliverable. Since the mailing piece didn't have to be identified as undeliverable at both attempts, there are sample addresses identified as by the USPS as undeliverable for which we received a mail response. As the last two columns of Table 3 show, the undeliverable rate was about 24 percentage points in the Puerto Rico test site. This rate is defined as the ratio of undeliverables to the mailout. The rate increases as the stratum get more rural in nature. The USPS was not able to deliver to about 47 percent of the sample addresses in the "Not MSA" stratum.

Table 3. Unmailable and Undeliverable Rates by Stratum

Stratum	Sample Size	Number Unmailable	Unmailable Rate and CI	Mailout	Number Undeliverable	Undeliverable Rate and CI
San Juan PMSA	6,277	365	5.8% ± 0.5%	5,912	1,048	17.7% ± 0.8%
Other MSA	3,878	423	10.9% ± 0.9%	3,455	985	28.5% ± 1.2%
Not MSA	1,845	256	13.9% ± 1.4%	1,589	745	46.9% ± 2.0%
Total*	510,854	41,665	8.2% ± 0.4%	469,189	110,974	23.7% ± 0.5%

*fully weighted sample

How do these rates for Puerto Rico compare to rates for states in the 2001 Supplementary Survey?

Table 4 below shows the comparison of the unmailable and undeliverable rates for the states in the 2001 Supplementary Survey compared to the Puerto Rico site totals. As the table shows, about 25 percent of the states had a higher unmailable rate and about 14 percent had a higher undeliverable rate. The remaining states had rates better than Puerto Rico. The table in Appendix 2 shows the detailed statistics for each state in the 2001 Supplementary Survey for the months of November and December 2001 combined. As the table in Appendix 2 shows, Arkansas (unmailable rate=6.3%, undeliverable rate=17.9%) and North Carolina (unmailable rate=6.9%, undeliverable rate=18.9%) have unmailable and undeliverable rates similar to the San Juan PMSA. There is no state that looks like the "Other MSA" stratum in both statistics; however, Mississippi comes closest (unmailable rate=10.2%, undeliverable rate=20.3%). The rates for North Dakota, (unmailable=13.9%, undeliverable=26.5%), Idaho (unmailable rate=15.8%, undeliverable rate=26.4%), and Alaska (unmailable rate=18.6%, undeliverable rate=25.7%) are most similar to the "Not MSA" stratum. No state had an undeliverable rate as high as the "Not MSA" stratum but several states had higher "unmailable rates" than all three stratum in this test. West Virginia had the highest unmailable rate; almost 34 percent of the addresses were unmailable.

Table 4. Comparison of Puerto Rico and Stateside Unmailable and Undeliverable Rates

	Unmailable Rate	Undeliverable Rate
Puerto Rico	8.2% ± 0.4%	23.9% ± 0.6%
Percent (and number) of States with Rates above PR	25.5% (13)	13.7% (7)
Percent (and number) of States with rates below PR Rate	74.5% (38)	86.3% (44)

B. Mail Response

What were the mail response rates? How did they vary by stratum?

Table 5 shows the mail response rates by stratum and for Puerto Rico. Mail response is defined as the number of returns received as a percentage of the mailout universe. As the table shows, the overall mail response rate in Puerto Rico was 27.8 percentage points across November and December. The total line represents the weighted total. Differences responses rates between the "San Juan MSA" and "Other MSA" stratum are not significant.

Table 5. Mail Response Rate by Stratum and for the Puerto Rico Test Site

Stratum	Mailout	Number Returned	Mail Response Rate ± Confidence Interval
San Juan MSA	5912	1723	29.1 % ± 1.0%
Other MSA	3455	959	27.8% ± 1.2%
Not MSA	1589	307	19.3% ± 1.6%
Total*	469,189	130,397	27.8% ± 0.7%

* fully weighted sample

How do these rates for Puerto Rico compare to rates for states in the 2001 Supplementary Survey?

As Table 6 below shows, no state in the 2001 Supplementary Survey had a mail response rate as low as Puerto Rico in November and December. Two thirds of the states had response rates between 50 and 60 percentage points. The table in Appendix 2 shows the combined state-level mail response rates for November and December, 2001 for each state in the 2001 Supplementary Survey. As these data show, the District of Columbia had the lowest mail response rate (40.8%.)

Table 6. Comparison of Puerto Rico and Stateside Mail Response Rates

Puerto Rico	27.8% ± 0.6%
Mail Response Rates	Number of States
less than 40%	0
40-50%	16
50-60%	30
>60%	5

How do the Puerto Rico Test Site mail response rates for the sample panels compare?

Table 7 shows the mail response rate for the November and December sample panels by stratum. The overall mail response rate for the Puerto Rico Test is similar in November and December; that is, about 26 percentage points. Looking at the stratum-level data shows the same trends seen in the combined results--that the metropolitan areas seem to have a higher mail response rate than the non-metropolitan areas.

Table 7. Mail Response Rates by Panel and Stratum

Stratum	November			December		
	Mailout	Number Returned	Mail Response Rate ± Confidence Interval	Mailout	Number Returned	Mail Response Rate ± Confidence Interval
San Juan PMSA	2,955	865	29.3% ± 1.3%	2957	858	29.0% ± 1.4%
Other MSA	1,731	460	26.6% ± 1.7%	1,724	499	28.9% ± 1.8%
Not MSA	788	135	17.1% ± 2.1%	801	172	21.5% ± 2.4%
Total*	234,496	64,079	27.3% ± 1.0%	234,693	66,317	28.3% ± 1.0%

* fully weighted sample

C. Contractor's Deliverability Flag

The contractor's file includes a variable called deliverability code (DLV). The code values are:

1: "Exact match" to the DSF address including ZIP+4.
2: "Equivocated match" to the DSF including ZIP+4. Some parts of the address do not match.
3: "Unmatched but deliverable"--The address doesn't match to the DSF but the full address is available
4: "Undeliverable"--The address is missing critical address information

To determine how effective the contractor's deliverability flag (DLV) was at identifying unmailable and undeliverable addresses, we considered two things:

How many addresses identified as undeliverable by the contractor did our mailability edit remove, and ***was the USPS able to deliver questionnaires to addresses considered mailable by our edit but undeliverable by the contractor (DLV=4)?***

Headquarters staff did not use the DLV code in the mailability edit to determine the universe of addresses for the mailout. Table 8 shows the proportion of addresses mailed by DLV code. As the table shows, the contractor identified a total of 2,274 addresses as undeliverable (DLV=4). The mailability edit removed only about 33 percent of these addresses. The mailability edit also removed about four percent of the addresses the contractor said were exact matches to the DSF, including Zip+4. The edit did not remove any addresses that were equivocated matches (DLV=2) or "Unmatched but deliverable" (DLV=3). The data in Table

8 seem to suggest that our edit rules did not catch all of the unmailable addresses. The stratum-level data are in Appendix 3. We see the same trends seen in the combined results–the mailability edit did not catch many addresses considered undeliverable by SEEK Data's edit.

Table 8. Unmailable Rate by Deliverability Code

DLV Code	Total Puerto Rico		
	Sample	Mailout	Unmailable Rate ± Confidence Interval
1	6,985	6,689	4.2% ± 0.4%
2	1,370	1,370	0.0% ± 0.0%
3	1,371	1,371	0.0% ± 0.0%
4	2,274	1,526	32.9% ± 1.7%

A total of 7,967 questionnaires were mailed out for which no mail return was checked in. Table 9 shows for these addresses the distribution of undeliverables by DLV code. The contractor considered addresses with DLV=1 and 2 as deliverable. The contractor considered addresses with DLV=3 to be potentially deliverable. The contractor said that the addresses with DLV=4 were not deliverable. As the data in Table 9 show, the explanation for the non-returns with DLV code of 3 or 4 was that the addresses were undeliverable. This is not the case for non-returns with a DLV code of 1 or 2. The stratum-level data are in Appendix 4. We see the same trends in these data.

Table 9. UAA Rate for Non-returns by DLV Code

DLV Code	Total Puerto Rico		
	Mailout without return	Number Undeliverable	Undeliverable Rate ± Confidence Interval
1	4,368	558	8.0% ± 0.5%
2	965	173	12.45% ± 1.4%
3	1,117	566	40.1% ± 2.2%
4	1,517	1,481	96.9% ± 0.7%

Another way to answer this question is to consider the number of responses we received from addresses mailed with a DLV code=4. This tells us what we might have lost if we had used the DLV code to determine housing unit mailout. As the data in Table 10 show, we received returns from less than one percent of the questionnaires (9) that the contractor said were undeliverable (DLV=4).

Table 10. Mail Response Rates for Stratum by Deliverability Code

DLV Code	Total Puerto Rico		
	Mailout	Number Returns	Mail Response Rate ± Confidence Interval
1	6,689	2,321	34.7% ± 1.0%
2	1,370	405	29.6% ± 2.0%
3	1,371	254	18.5% ± 1.7%
4	1,526	9	0.6% ± 0.3%

Based on these data, it appears that the contractor's DLV code is a good indicator of deliverability by the USPS. About 19 percent of the addresses in our sample were coded as DLV=4. Of these, we mailed to 1,526 addresses. The USPS said that about 94 percent (1,433) of these were undeliverable when they attempted to deliver. Of the 93 questionnaires delivered, we received returns from only nine addresses. Based on these results, we should consider excluding addresses where DLV=4 from future samples in Puerto Rico.

D. Cooperation

How cooperative were respondents who received a questionnaire?

Typically we calculate mail return rates as the ratio of the number of mail returns to the number of occupied housing units. This is considered a better measure of respondent cooperation to a survey or census than mail response rates. This test, however, did not follow up to determine the universe of occupied housing units. An alternative estimate of respondent cooperation for Puerto Rico can be obtained by looking at how many returns we received compared to the number of addresses in sample that received a questionnaire; that is, the ratio of returns to the mailout universe less the undeliverables and the DLV=4 cases for which we did not receive a return. Table 11 shows the cooperation rate by stratum. Of the 510,854 sample addresses, theoretically about 68 percent of them received a questionnaire and should have been eligible to respond; that is, they were not considered unmailable or undeliverable by the USPS or SEEK Data.

As the data in Table 11 show, we received a completed questionnaire from 36 percent of all addresses in the Puerto Rico Test that we estimate received a form in the mail. This is consistent by stratum and panel. One possible explanation for these rates is vacancy. Vacants were not identified and therefore are included in the denominator unless the USPS identified them as undeliverables.

Table 11. Cooperation Rate for Puerto Rico and by Stratum

Stratum	Sample Size	Mailed and not Undeliverable	Percentage	Mail Returns	Cooperation Rate ± Confidence Interval

San Juan	6,277	4,885	77.8	1,723	35.3% ± 1.0%
Other MSA	3,878	2,485	64.1	959	38.6% ± 1.3%
Not MSA	1,845	852	46.2	307	36.0% ± 1.8%
Total	510,854*	360,056	70.5	130,397	36.2% ± 0.6%

* fully weighted sample

How do these rates for Puerto Rico compare to rates for states in the ACS?

Appendix 2 shows the November-December cooperation rates for the states in the 2001 Supplementary Survey. Since the DLV variable was not a stateside variable, we calculated the cooperation rate by subtracting the undeliverables for which we did not get a return, from the mailout universe. As the data in Appendix 2 show, the state-level cooperation rates range from 49% for New York to almost 73% for Vermont. Table 12 below shows a summary of the stateside cooperation rates compared to Puerto Rico. As the data show, every state had a higher cooperation rate than Puerto Rico. The majority of states had a cooperation rate of at least 50 percentage points.

Table 12. Comparison of Puerto Rico and Stateside Cooperation Rates

Puerto Rico	36.2% ± 0.6%
Cooperation Rate	Number of States
less than 40%	0
40-50%	2
50-60%	19
60-70%	23
> 70%	7

E. Cost Modeling

What do these data tell us about cost modeling assumptions for Puerto Rico in 2003?

Tables 13a and 13b provide cost estimates for two simulations for the 2003 American Community Survey.

1. 36 percent mail response and 20 percent unmailable rate in Puerto Rico[3]
2. No mail data collection in Puerto Rico

It assumes:

- Total addresses 1,467,250
  Census 2000 1,327,250
  Contractor Adds 140,000
- Sampling Rate 2.50%
- Sample Size 36,681
- Mail Cost per Case $15[4]
- CAPI Cost per Case $150
- Unmailable cases are subsampled at 2 in 3 for CAPI
- No cases are completed by CATI

As the data in Tables 13a and 13b show, conducting mail data collection in Puerto Rico, under the specified assumptions, will cost about $1.6 million less than if no mail data collection is used in Puerto Rico.

---

[3] This assumes an unmailable rate based on the DLV variable and that the mail response rate will be higher in 2003 because we will exclude unmailable addresses (using the DLV variable) from the sampling universe.

[4] These estimates are based on a rough assumption for cost per case for the different modes. If we send cases to CATI, this would reduce the CAPI work load. The costs associated with the CATI operation would reduce the CAPI work load and costs. Therefore, we estimate that the tottal operational costs would be less than the total cost estimate shown here.

Table 13a. Cost Estimate for Puerto Rico assuming 36% Mail Response and 20% Unmailable Rates

Assumptions		Estimated Cost
Sample Size	36,681	
Mail Response Rate	36%	
Unmailable Rate	20%	
Mailed Forms	29,345	$440,175
Returned Forms	10,564	
Unreturned Forms	18,781	
Unmailable Forms	7,336	
CAPI from Mailable Universe (1/3 of unreturned)	6,260	$939,000
CAPI from Unmailable Universe (2/3 of unmailables)	4,891	$733,650
CAPI Total	11,151	$1,672,650
Total Cost		$2,112,825

Table 13b. Cost Estimate for Puerto Rico assuming No Mail Data Collection

Assumptions		Estimated Cost
Sample Size	36,681	
Mail Response Rate	0%	
Unmailable Rate	100%	
Mailed Forms	0	$0
Returned Forms	0	
Unreturned Forms	0	
Unmailable Forms	36,681	
CAPI from Mailable Universe (1/3 of unreturned)	0	$0
CAPI from Unmailable Universe (2/3 of unmailables)	24,454	$3,668,100
CAPI Total	24,454	$3,668,100
Total Cost		$3,668,100

5. Conclusions

This test was successful because we were able to mail questionnaires to and get responses from more than a 27 percent of the addresses in this test. This is a good first start given that this is the first time the Census Bureau has attempted to mail to addresses in Puerto Rico to complete census or survey data collection. Also, as the cost data show, collecting data by mail could save over $1.6 million, assuming a 36 percent mail response rate and 20 percent unmailable rate.

While the preliminary indications show promise, there are still problems and work to be done. The undeliverable and unmailable rates are high in Puerto Rico. As the data in Appendix 2 show, there are several states that had higher undeliverable and unmailable rates in November and December, 2001 than Puerto Rico. However, no state had mail response or cooperation rates as low as Puerto Rico.

The data presented here suggest that more work is needed to ensure that we have an improved MAF for Puerto Rico before the 2003 full implementation. We should consider using SEEK Data's DLV variable in defining the universe of unmailable addresses.

We have not learned all we need to about conducting mail data collection in Puerto Rico. As mentioned in the limitations, our results assume that mail returns were completed by the correct sample address. In this test we had no way of verifying this. We will not understand how well this worked until we attempt to contact these units to resolve edit failures.

Appendix 1

Variables used in Analysis by Source

Source	Variable	Description	Code	Code Description
Control File	PANEL	PR Mailout Panel	200111	November panel
			200112	December panel
	M1_ST and	Mailback codes for	01	Form not mailed
	M2_ST	1st and 2nd mailings	10	Form not returned
			all others	Form returned
PMR File	M01REASN and	Undeliverable codes for 1st	0	Not a Undeliverable
	M02REASN	and 2nd mailings	1	Bad Address
			2	Unkown Address
			3	Vacant
			4	Refused
			5	Other
Contractor's File	STRATA	MSA Stratum code	1	San Juan PMSA
			2	Other MSA area
			3	Not in an MSA
	DLV	DLV Code	1	""Exact match""
			2	""Equivocated match""
			3	""Unmatched but deliverable""
			4	""Undeliverable""

Appendix 2

State-level Rates Related to Mailability for the November and December 2001 Panels

State Abbrev.	Sample Size	Not Mailed	Mail Universe	Not UAA Returned	Not UAA Not Ret.	UAA Returned	UAA Not Ret.	Mail Rate	Return Rate	UAA Rate	Cooperation Rate
AL	1783	132	1651	781	609	14	247	92.60%	48.15%	15.81%	56.62%
AK	1123	209	914	380	299	3	232	81.39%	41.90%	25.71%	56.16%
AZ	4457	119	4338	2326	1432	32	548	97.33%	54.36%	13.37%	62.22%
AR	1328	83	1245	560	462	8	215	93.75%	45.62%	17.91%	55.15%
CA	13245	172	13073	6300	5560	45	1168	98.70%	48.54%	9.28%	53.30%
CO	1593	15	1578	876	529	9	164	99.06%	56.08%	10.96%	62.59%
CT	1256	6	1250	672	468	4	106	99.52%	54.08%	8.80%	59.09%
DC	1073	0	1073	434	523	4	112	100.00%	40.82%	10.81%	45.58%
DE	1082	41	1041	533	355	9	144	96.21%	52.07%	14.70%	60.42%
FL	9229	95	9134	4427	3439	91	1177	98.97%	49.46%	13.88%	56.78%
GA	3236	206	3030	1383	1186	14	447	93.63%	46.11%	15.21%	54.08%
HI	1085	63	1022	518	361	4	139	94.19%	51.08%	13.99%	59.12%
ID	972	154	818	418	184	11	205	84.16%	52.44%	26.41%	69.98%
IL	5540	74	5466	2944	2013	34	475	98.66%	54.48%	9.31%	59.67%
IN	2460	40	2420	1375	746	13	286	98.37%	57.36%	12.36%	65.04%
IA	2325	46	2279	1367	573	18	321	98.02%	60.77%	14.87%	70.74%
KS	1586	78	1508	860	407	15	226	95.08%	58.02%	15.98%	68.25%
KY	2557	278	2279	1168	725	19	367	89.13%	52.08%	16.94%	62.08%
LA	2403	108	2295	964	863	17	451	95.51%	42.75%	20.39%	53.20%
ME	988	221	767	384	188	13	182	77.63%	51.76%	25.42%	67.86%
MD	2431	53	2378	1273	812	17	276	97.82%	54.25%	12.32%	61.37%
MA	3670	14	3656	2121	1257	12	266	99.62%	58.34%	7.60%	62.92%
MI	3840	90	3750	2141	1151	19	439	97.66%	57.60%	12.21%	65.24%
MN	1799	44	1755	1134	444	15	162	97.55%	65.47%	10.09%	72.13%
MS	2500	254	2246	916	874	14	442	89.84%	41.41%	20.30%	51.55%
MO	2443	215	2228	1267	650	12	299	91.20%	57.41%	13.96%	66.30%

State Abbrev.	Sample Size	Not Mailed	Mail Universe	Not UAA Returned	Not UAA Not Ret.	UAA Returned	UAA Not Ret.	Mail Rate	Return Rate	UAA Rate	Cooperation Rate
MT	1492	292	1200	619	261	9	311	80.43%	52.33%	26.67%	70.64%
NE	2567	97	2470	1444	691	29	306	96.22%	59.64%	13.56%	68.07%
NV	1048	33	1015	463	371	6	175	96.85%	46.21%	17.83%	55.83%
NH	986	102	884	500	238	10	136	89.66%	57.69%	16.52%	68.18%
NJ	3009	9	3000	1617	1165	12	206	99.70%	54.30%	7.27%	58.30%
NM	1286	205	1081	472	356	4	249	84.06%	44.03%	23.40%	57.21%
NY	9671	209	9462	4280	4430	32	720	97.84%	45.57%	7.95%	49.33%
NC	3168	219	2949	1399	993	11	546	93.09%	47.81%	18.89%	58.68%
ND	1044	145	899	473	202	16	208	86.11%	54.39%	24.92%	70.77%
OH	6201	54	6147	3592	1940	33	582	99.13%	58.97%	10.00%	65.14%
OK	1401	149	1252	626	418	11	197	89.36%	50.88%	16.61%	60.38%
OR	3354	7	3347	1993	1057	23	274	99.79%	60.23%	8.87%	65.60%
PA	5462	349	5113	3053	1516	26	518	93.61%	60.22%	10.64%	67.01%
RI	1080	7	1073	586	403	5	79	99.35%	55.08%	7.83%	59.46%
SC	1587	121	1466	695	536	10	225	92.38%	48.09%	16.03%	56.81%
SD	1394	108	1286	784	300	9	193	92.25%	61.66%	15.71%	72.55%
TN	2465	91	2374	1200	875	17	282	96.31%	51.26%	12.59%	58.17%
TX	8711	541	8170	3701	3387	39	1043	93.79%	45.78%	13.24%	52.48%
UT	1027	36	991	516	295	19	161	96.49%	53.99%	18.16%	64.46%
VT	1017	218	799	410	159	16	214	78.56%	53.32%	28.79%	72.82%
VA	2803	116	2687	1471	912	20	284	95.86%	55.49%	11.31%	62.05%
WA	2819	25	2794	1573	827	15	379	99.11%	56.84%	14.10%	65.76%
WV	2084	705	1379	734	403	11	231	66.17%	54.02%	17.55%	64.90%
WI	2549	34	2515	1434	553	20	508	98.67%	57.81%	20.99%	72.45%
WY	954	73	881	410	186	7	278	92.35%	47.33%	32.35%	69.15%

Table 8. Unmailable Rate by Stratum and Deliverability Code

DLV Code	Total Puerto Rico			San Juan			Other MSA			Not MSA		
	Sample	Mailout	Unmailable Rate± Confidence Interval	Sample	Mailout	Unmailable Rate ± Confidence Interval	Sample	Mailout	Unmailable Rate± Confidence Interval	Sample	Mailout	Unmailable Rate ± Confidence Interval
1	6985	6689	4.24% ± 0.4%	3658	3576	2.2% ± 0.4%	2406	2299	4.5% ± 0.7%	921	814	11.6% ± 1.8%
2	1370	1370	0.00% ± 0.0%	1121	1121	0.0% ± 0.0%	202	202	0.0% ± 0.0%	47	47	0.0% ± 0.0%
3	1371	1371	0.0% ± 0.0%	775	775	0.0% ± 0.0%	429	429	0.0% ± 0.0%	167	167	0.0% ± 0.0%
4	2274	1526	32.9% ± 1.7%	723	440	39.1% ± 3.0%	841	525	37.6% ± 3.5%	710	561	21.0% ± 2.8%

Table 9. Distribution of Undeliverables by DLV Code

DLV Code	Total Puerto Rico			San Juan			Other MSA			Not MSA		
	Mailout without return	Undeliver-able	Undeliver-able Rate ± CI	Mailout without return	Undeliver-able	Undeliver-able Rate ± CI	Mailout without return	Undeliver-able	Undeliver-able Rate ± CI	Mailout without return	Undeliver-able	Undeliver-able Rate ±CI
1	4368	558	8.0% ± 0.5%	2351	226	6.3% ± 0.7%	1473	228	9.9% ± 1.0%	544	104	12.8% ± 1.9%
2	965	173	12.5% ± 1.4%	782	131	11.7% ± 1.6%	148	35	17.3% ± 4.4%	35	7	14.9% ± 8.6%
3	1117	566	40.1% ± 2.2%	618	268	34.6% ± 2.8%	356	211	49.2% ± 4.0%	143	87	52.1% ± 6.4%
4	1517	1481	96.9% ± 0.7%	438	423	96.1% ± 1.5%	519	511	97.3% ± 1.2%	560	547	97.5% ± 1.1%