

Determining the Effectiveness of Multiple Nonresponse Followup Contact Attempts on Response and Data Quality¹

Jennifer Guarino Tancreto and Michael Bentley

U.S. Census Bureau, 4700 Silver Hill Road, Stop 9200, Washington, DC 20233

Abstract

Nonresponse Followup is one of the most expensive operations in a decennial census. The number of contact attempts that enumerators make to collect data for non-responding households is one area that could be considered in an effort to cut back costs. Official Census 2000 enumeration procedures specified that field staff should make up to six contact attempts to collect data from households that failed to return their census forms. We implemented an ad hoc research study to determine if we could feasibly use fewer than six contact attempts to collect data from these households without impacting response or data quality. The results suggest that enumerators rely on the sixth contact to make last chance efforts to find a respondent, resulting in a significant increase in the interview rate at contact 6 compared to contact 5. In this sense, the sixth contact appears to be critical in achieving high response; however, we believe that this gain could be achieved at an earlier contact that is set as a limit.

1. Background

Contacting households that fail to return a census form by mail is a very costly operation in a decennial census. The total cost of this operation, called Nonresponse Followup (NRFU), in Census 2000 was 1.1 billion dollars, with an approximate \$26.96 cost per case (Moul, 2002). Given these sizeable costs, the Census Bureau is actively searching for means of reducing these costs without jeopardizing the quality of data.

One area that could be considered in the effort to cut back costs is the number of visits that field staff make to collect data for non-responding households. It is likely that costs will decrease if the maximum number of enumerator visits

decreases, but there are tradeoffs that need to be considered. For instance, fewer contact attempts could result in decreased data quality, due to a potential increase in the number of proxy responses.

Official Census 2000 Nonresponse Followup enumeration procedures specify that field staff should make up to six contact attempts (3 personal visit, 3 telephone) to collect data from non-responding households (Burt *et. al.*, 2003). Currently there are no empirical census data to support this determination, but the rationale is to attempt contact with a household member (rather than a proxy) to provide better coverage and data quality.

It seems logical that repeated contact attempts with attention to time scheduling will produce higher response. There are many studies on telephone interviewing that show an increasing trend in response after multiple call attempts (Groves *et. al.*, 1998). However, many studies indicate that there is a point where additional contacts result in diminishing returns. Bates (2003) found that the first through fifth contact attempts in the Survey of Income and Program Participation resulted in an above-average percentage of interviews; attempts made after the fifth attempt had a below-average chance of resulting in an interview.

To date, there has been no research that studied the outcome of contact attempts in a decennial census. The purpose of this study was to determine if the Census Bureau could feasibly use fewer contacts to achieve the same high quality of data. Specifically, we set out to determine whether repeated contacts during NRFU increased the likelihood of response, especially household member response, to the census. Additionally, we studied the relationship of repeated contacts to data quality, as measured by form completeness. Finally, we analyzed the Census 2000 contact data by time and day of attempt to

¹This report is released to inform interested parties of research and to encourage discussion. The views expressed on statistical, methodological, technical, or operational issues are those of the authors and not necessarily those of the U.S. Census Bureau.

understand whether enumerator visits can be further optimized.

Our hypotheses were as follows:

1. There is a point where repeated contact attempts fail to produce a significant number of interviews.
2. The larger the number of contact attempts, the more likely the data are provided by a proxy respondent. This hypothesis stems from Census 2000 enumerator procedures, which dictate that enumerators should seek proxy respondents after making six (3 personal visit, 3 telephone) attempts to contact NRFU households (or after 3 personal visits for units with no available telephone number).
3. The more contact attempts made by an enumerator, the more complete data obtained. This hypothesis assumes that more contacts may lead to interaction with different household members or that more contacts may be required if time was limited at previous visits.
4. Enumerators attempt contact more frequently on weekdays than weekends. Hence, more interviews are conducted on weekdays than weekends.
5. Contact attempts result in interviews more frequently during a weekday evening compared to weekends and other times during the weekdays.

2. Methodology

This research study used the Master Trace Sample (MTS) database as its data source. The MTS is a nationally representative systematic sample of housing units in Census 2000. The MTS database links enumerator contact data with response data for research purposes.

The universe for this study consisted of cases that did not return a census questionnaire by a specific date. We refer to this population as the NRFU universe. We restricted the analysis to short form cases so that the results are applicable to the methodology planned for the 2010 Census². Moreover, cases included in special enumeration operations were eliminated because enumeration procedures differed for these housing units. The final analytical universe contained 697,789 NRFU contacts for 319,656 different housing units.

² Census 2000 included a short questionnaire that collected basic demographics and a long form that requested more extensive information. The 2010 Census will include the short questionnaire only, as the information collected in the longer questionnaire is now being collected in the American Community Survey.

2.1 Analysis variables and data considerations

Contact data: Contact data, including date, time, mode, and outcome of each contact attempt, were collected on the front cover of the enumerator questionnaires. The NRFU operation used a paper and pencil data collection method.

Form completeness: We used the amount of completed data among the person items (name, age, date of birth, sex, race, Hispanic Origin) to indicate data quality.

Seasonal areas: We flagged housing units in areas with at least 20 units where at least 40 percent were seasonal vacant units so we could determine whether these areas exhibit different contact outcomes than other areas.

Enumerator characteristics: We eliminated cases enumerated by crew leaders, since crew leaders may have different characteristics compared to regular enumerators. We identified crew leaders as those enumerators who did not indicate that they spoke a language other than English and who worked a low NRFU caseload relative to the average (by Local Census Office). We defined low NRFU caseload outliers as values which were less than or equal to 0.6 times the quantity of the first quartile minus the interquartile range. We eliminated approximately 1.3 percent of all contacts in the database using this methodology, as we suspect that crew leader assistants may have conducted these cases.

Similarly, we also removed cases worked by enumerators with unusually high NRFU caseloads by targeting values that were greater than 3 times the quantity of the third quartile plus the interquartile range. We felt that extremely high caseloads were likely recorded in error. This removed an additional 1.6 percent of the contacts in the MTS database.

2.2 Statistical analyses

2.2.1 Table comparisons

We first examined cross tabulations of the number of contact attempts and contact outcome, contact mode, respondent type (i.e. household member or proxy), and form completeness. For the tables that present data by contact number, we compared the value of the variables of interest across contacts 3 through 6 to determine if the variables differed among later contacts (i.e., 4, 5 and 6 contacts) as compared to earlier (i.e., 3) contacts. Using Bonferroni's Multiple Comparison procedure, we assured that the family-wise error rate across the various comparisons did not exceed 10 percent.

2.2.2 Logistic regression models

We implemented modeling to control for factors that may confound any relationship between the variables of interest. These factors included enumerator experience, characteristics (such as education and knowledge of a language other than English), workload and geographic spread³. Also, we controlled for the mode of the contact attempt (personal visit or telephone) and any differences from regional management by including regional office in the models. Other predictor variables included an indicator of seasonal areas, and enumerator test scores that are used in hiring enumerators.

The regression models helped determine whether any relationships in the tables remained in the presence of confounding factors. Because our sample was very large, many factors showed statistical significance without exhibiting a substantial effect. Therefore, we used statistical significance as well as odds ratios to determine which factors had a substantive impact on the dependent variables.

3. Limitations

Quality of contact data: Anecdotal reports suggested that enumerator training did not emphasize the importance of collecting contact data relative to the other data on the form. As a result, the data are often unreliable. Also, the contact outcome codes were not mutually exclusive and did not distinguish partial from completed interviews. Enumerators were instructed to enter one of the following outcomes for each attempt: no contact, conducted interview, left notice of visit, refusal, other.

Enumerator characteristics: It was possible for multiple enumerators to work a case in Census 2000. We used the characteristics of the last enumerator who worked the case.

Potential confounding variables: The Census 2000 Partnership and Marketing Program established community connections in some areas to help gain trust. There are no variables on the MTS database that identify the areas in which these arrangements may have been in place, and consequently, no way to determine possible confounding effects.

No cost-benefit analysis: Cost data are not available on the database. Therefore, we could not analyze any potential costs savings from fewer enumerator visits.

3 We used correlations to remove predictor variables that were highly correlated with other predictors in the models. We also removed those that had highly skewed binomial distributions.

Study design: This report documents the results of an ad hoc study without an experimental design. The results describe what occurred in Census 2000, but do not permit inferences.

4. Results

4.1 Contact outcome

We first present the cumulative interview rate as well as the gain in the interview rate from the previous contact in Table 1. Note that the figures in Table 1 reflect information from each household's last contact attempt. For example, we only used data from the third contact attempt when three contact attempts were made.

All together, enumerators recorded "conducted interview" for 85 percent of housing units⁴. About 41 percent were interviewed after just one attempt, 61 percent after two attempts, and 71 percent after three attempts. After this, the increase in the interview rate was more gradual. The figure increased to 77 percent after four attempts, then 80 percent after five contacts.

The marginal gain in the cumulative interview rate decreased with each contact, with the exception of a modest increase at contact 6 (all differences significant at $p < .0001$). This exception from the trend at contact 6 is likely due to enumerators' last chance efforts to interview proxy respondents.

Table 1. Cumulative Total and Percentage Point Increase for "Conducted Interviews"

Contact Number	Interviews	Cumulative Percent	Percentage Point Increase
	Conducted at Each Contact		
1	131,495	41.1	41.1
2	62,367	60.6	19.5
3	32,996	71.0	10.3
4	18,619	76.8	5.8
5	11,295	80.3	3.5
6	16,070	85.4	5.0

Overall, enumerators averaged 2.1 contact attempts with housing units that were interviewed on their final contact. In comparison, 3.1 attempts were averaged for households with a final outcome of refusal.

Next, Table 2 provides a picture of the outcomes at each contact. Note that this table includes multiple contacts for households where more than one contact was attempted.

4 We could not determine final contact outcome for all cases; therefore, it is possible that not all of the remaining 15 percent were non-interviews at their last contact.

For example, this table includes data for contacts 1, 2 and 3 for a household that required three contact attempts.

The table shows that 42.1 percent of the initial contacts were recorded as interviews⁵. As expected, the rate of successfully conducting interviews dropped from about 42 percent at first contact to 33 percent at the fifth contact, before significantly increasing to almost 71 percent at the sixth and final contact⁶. Again, we believe that enumerators' last chance efforts to interview proxy respondents are driving this outcome for contact 6.

Table 2. Distribution of Outcomes at Each Contact⁷

Contact Number	Inter-view	Contact Outcome (Percent)			
		Left Notice of Visit	Refusal	No Contact	Other ⁸
1	42.1	35.2	1.3	8.2	13.1
2	38.5	18.9	1.6	28.6	12.4
3	35.7	18.7	1.9	30.8	13.1
4	33.2	17.3	1.9	31.3	16.4
5	32.6	16.2	2.3	32.5	16.3
6	70.6	3.3	1.9	8.3	16.0
Total	40.1	25.6	1.6	19.3	13.5

We also studied the distribution of contact outcomes by mode of contact (personal visit or telephone) (tables not shown). Since 94 percent of attempts were made by personal visit, the outcomes for personal visit contacts were very similar to Table 2. The distribution for telephone contacts may indicate that enumerators sometimes have more success in obtaining interviews over the phone. Almost 54 percent of telephone contacts resulted in interviews on the second attempt (all first attempts were personal visits) compared to 35 percent of personal visits. It is possible that this resulted from enumerators not always recording contact attempts for phone calls that went unanswered. Also, households with available telephone numbers may have been more cooperative than those with unlisted numbers.

Next, we ran a logistic regression model to study the effect of the number of contacts and other predictor variables⁹ on

final contact outcome (interview or no interview). The results support the findings in Tables 1 and 2. The chances of conducting an interview diminished as the number of contact attempts increased. The odds of obtaining an interview dropped by about 13 percent with each successive contact attempt. Note that the increase in the interview rate at contact 6 shown in Tables 1 and 2 was offset by the decreases between contacts 1 and 5 in the model. Another interesting result from the model was that the odds of an interview at the last contact were about 3.5 times higher when the final contact was made with a household member rather than a proxy.

4.2 Respondent type

Table 3 lists the distribution of respondent type (household member or proxy respondent) for each final contact recorded as "conducted interview." Overall, just under two-thirds (63.6 percent) of interviews had a household member respondent. This percent was similar between contacts 1 through 5, before decreasing significantly at contact 6 to just below half. This decrease in household respondents for contact 6 coincided with a significant increase in proxy respondents. We anticipated this shift, since enumerators often resort to proxies to collect data for households they were unable to contact. This increase in proxy reporting is important to note, since proxy responses typically have poorer data quality than household responses, as we show later in this report.

Table 3. Distribution of Respondent Type at Last Contact for "Conducted Interviews"

Contact Number	Number of Cases	Household Respondent (%)	Proxy Respondent (%)	Missing (%)
1	131,495	64.4	33.4	2.2
2	62,367	66.1	32.0	1.9
3	32,996	65.4	32.6	2.1
4	18,619	62.2	36.0	1.8
5	11,295	59.7	38.3	2.0
6	16,070	47.8	50.4	1.8
Total	272,842	63.6	34.4	2.0

The second logistic regression model studied the effects of the number of contacts and other predictor variables on respondent type (household member or proxy). The results concurred with the findings in Table 3. That is, holding other variables constant, the odds of interviewing a household member were about 9 percent lower than the odds of interviewing a proxy with each successive contact.

We also found that the odds of interviewing a household member were about 5 percent less for personal visit

distance, regional office, and seasonal area.

5 Note that this figure differs slightly from Table 1. We found that some cases with a "conducted interview" status at contact 1 received additional contacts. Therefore, they are included in Table 1 at their last contact number.

6 The drops are statistically significant from the third to the fourth and fifth contacts, with $p < .0001$.

7 The figures may not add up to 100 as expected due to the rounding.

8 Includes missing, undetermined, and other outcomes.

9 Variables included: proxy/household respondent, contact mode, enumerator education/experience/language/travel

contacts compared to telephone contacts. In high seasonal vacant areas, the odds of interviewing a household member were 95 percent lower compared to other areas.

4.3 Form completeness

Table 4 examines average form completeness at final contact by respondent type for interviewed households. Clearly, returns with household member respondents had a higher average form completeness rate than proxy respondents. On average, household member reports were 98 percent complete compared to 86 percent for proxy reports ($p < .0001$). This is one concern about reducing the number of enumerator contacts; item nonresponse rates may suffer as a result of more proxy reporting.

There was a slight decreasing trend in average form completeness from contacts 1 to 6 for both household member and proxy respondents. The only significant differences in average form completeness were between contacts 1, 2 and 3 compared to contact 6 (p -values for both household member and proxy respondents were $< .0001$).

Table 4. Average Form Completeness by Respondent Type at Last Contact¹⁰

Contact Number	Household Member		Proxy		Overall Avg Form Comp (%)
	%	Avg Form Comp (%)	%	Avg Form Comp (%)	
1	65.9	98.5	34.2	88.7	97.7
2	67.4	98.4	32.6	86.6	97.1
3	66.7	98.1	33.3	85.6	96.2
4	63.4	98.0	36.7	86.2	95.4
5	61.0	98.0	39.1	84.8	94.6
6	48.7	97.2	51.3	83.9	91.6
Total	64.9	98.3	35.1	86.3	96.7

The third logistic regression model provided the effect of the variables on form completeness (greater than or equal to 96% complete, or less). The results indicated a slight negative association (about 8 percent lower odds for each successive contact) between the number of contacts and high form completeness. We also found that respondent type was a stronger predictor of form completeness than the number of contacts. The odds of obtaining highly complete interviews were more than 7 times higher for household respondents compared to proxies. Moreover, households in areas with high seasonal vacancy were less

likely to have high form completeness than households in other areas.

We reran the model with two interaction terms: respondent type and number of contacts, and respondent type and seasonal vacant area. We used this model to determine if the drop in form completeness due to multiple contacts and due to interviews in seasonal areas varies between household members and proxies. The results showed that the decrease in the form completeness by number of contacts was stronger for proxies than for household members. Furthermore, the drop in form completeness in seasonal areas was more pronounced for proxies than household members.

4.4 Timing of contact attempts

In Table 5, we studied the times and days when enumerators attempted contact and when interviews were conducted. For example, 13.7 percent of attempted contacts took place on a weekday morning, and 12.9 percent of cases that were interviewed took place on a weekday morning. Mornings were defined as 7:00 am to 11:59 am, afternoons as 12:00 pm to 4:59 pm, and evenings as 5:00 pm to 9:59 pm. All recorded times outside of these ranges were removed.

The distribution of days and times when interviews were conducted was very close to the distribution of attempted contacts, as we would expect since there must be an attempt to get an interview. We would expect differences if the interview rate showed greater success for a particular time and day combination than other times. Table 5 shows that 35 percent of conducted interviews occurred on weekday evenings, while 33 percent of all contacts occurred at that time. This difference is small, but it suggests that weekday evenings were particularly fruitful for making contact, and may have been slightly underutilized. In fact, we found that 44 percent of attempts on weekday evenings resulted in interviews, which was significantly higher than the rates for all other times (table not shown).

Table 5 also shows that enumerators favored weekdays over weekends. On weekdays, enumerators made slightly fewer attempts in the afternoons compared to the evenings ($p < .0001$). On weekends, fewer contacts attempts were made on Sundays than Saturdays, with both days showing more attempts in the afternoon than the morning or evening.

10 Cases with missing data for respondent type are included in Overall Average Form Completeness and excluded from the average form completeness for each respondent type.

Table 5. Percent Distribution of Contact Attempts and Interviews by Time and Day of the Week

Time of Day	Attempted Contacts			Interviews		
	M-F	Sat	Sun	M-F	Sat	Sun
Morning	13.7	3.9	1.1	12.9	3.8	1.1
Afternoon	30.3	7.6	4.5	29.8	7.3	4.3
Evening	33.3	3.0	2.5	35.3	3.0	2.5
Total	77.3	14.5	8.1	78.0	14.1	7.9

Finally, we studied contact times by contact number in Table 6. We observed few differences in the timing of contact attempts across contacts. One finding was that 35 percent of first attempts occurred on weekday afternoons and 29 percent on weekday evenings. These numbers flipped with the second contact, and continued through the remaining weekday contacts. It seems that many enumerators attempted initial contact during the day. After one unsuccessful afternoon attempt, they made repeated attempts on weekday evenings. Weekend attempts did not increase after a few previous attempts, despite the fact that the Enumerator Manual specifies “the best times to call are late afternoon and evenings until 9pm and during the day Saturday and Sunday” (Burt *et. al.*, 2003).

Table 6. Time and Day of Attempted Contact by Contact Number for All Contacts

Contact Number	Attempted Contact (%)					
	Weekday			Weekend		
	morn	aftr-noon	even-ing	morn	aftr-noon	even-ing
1	14.4	35.1	28.5	4.5	12.6	4.9
2	12.5	27.8	36.3	5.2	12.5	5.7
3	12.9	25.6	38.4	5.5	11.2	6.4
4	13.3	24.1	39.2	5.9	11.3	6.2
5	14.0	24.1	38.9	5.7	10.7	6.6
6	14.4	25.0	37.6	6.2	9.9	7.0
Total	13.7	30.3	33.3	5.0	12.1	5.6

5. Discussion

Please note that this is an ad-hoc research study with no experimental design. In Census 2000, there was a limit of six contacts per case that likely resulted in a substantial increase in the percentage of interviews conducted at the sixth contact attempt. If the limit were five contacts, we would likely see a large increase in interviews at the fifth contact rather than the sixth contact. As such, we cannot draw any conclusions about the effectiveness of “additional” contacts. We can only describe the results of, and speculate about, what happened.

How was the interview rate affected by each successive contact attempt?

As predicted, the rate of conducting interviews decreased with each consecutive contact attempt up to the fifth contact. As a result, the gain in overall response was only about 3.5 percent by the fifth contact. There was a large increase, however, in the interview rate from contact 5 (33 percent) to contact 6 (71 percent). In fact, 5 percentage points of overall response was gained at contact 6. This jump was almost certainly a result of last chance efforts to interview proxies at contact 6. It is likely that this gain could be observed at earlier contacts if less than six attempts per case were permitted.

How did the number of contacts affect the rate of proxy interviews?

Our results suggest that the chances of collecting data from a proxy held fairly steady in the 32 to 38 percent range for cases interviewed during contacts 1 through 5. Cases that required six contacts had a significantly higher proxy rate (50 percent) than earlier contacts. Similar to previous findings, we believe that this result stemmed from last resort efforts to obtain an interview at the final contact. We believe this trend would occur at any contact number that is set as the limit.

How did the number of contacts and enumeration characteristics affect data quality?

Overall, we found that interviews with household members had higher average form completeness (98 percent) than proxy interviews (86 percent). Contrary to our hypothesis, we found that form completeness decreased slightly as the number of contacts increased for both household and proxy interviews. In retrospect, this finding seems logical for two reasons. First, we found that the proxy rate increased at later contacts and that proxy interviews had less complete data than household interviews. Second, the number of contact attempts required to obtain an interview may indicate the general cooperativeness of a household or the difficulty associated with collecting its data.

Ultimately, the results indicated that respondent type was a more important determinant of form completeness than number of contacts. In fact, decreases in form completeness after repeated contact attempts were more pronounced for proxies than for household members.

When did enumerators attempt contact and when were interviews conducted?

The majority of contact attempts occurred on weekdays. The distribution of interviews followed closely the distribution of contact attempts, indicating that

enumerators optimized their visits well in Census 2000. The slight increase between the percent of contacts attempted on weekday evenings and the percent of interviews conducted during that time suggests that weekday evenings may be one of the best times to call.

6. Recommendations

This report describes the outcome of the contact attempts in Census 2000. Because enumerators were limited to six contact attempts per case, the results suggest that they used the sixth attempt to collect data from any source. Therefore, it is difficult to estimate what the impact to response and data quality would have been if the number of contacts were limited to five. It is unlikely that this question will ever be answered with the results of an ad-hoc study.

We recommend designing a split panel experiment, where random groups of enumerators are given different limits on the number of attempts permitted. With this experimental design, we can assess the effect of fewer attempts on response and data quality. This experiment could also include a cost-benefit analysis to judge the tradeoffs between data quality and costs.

We recognize that there may be significant operational challenges in designing this experiment. It may be challenging to ensure that areas with different contact limits have similar qualities. Additionally, this experiment might be difficult to implement due to training and management issues as well as the inability to move enumerators based on work needs.

Finally, we suggest allocating more resources to enhancing the quality of the record of contact data. Developments in this direction are already underway at the Census Bureau. Enumerators will use Hand Held Computing Devices for NRFU in the 2010 Census. These devices provide automated time and date stamps for each contact attempt. Moreover, the automated instrument includes edits to prompt enumerators that attempt to enter an invalid or blank contact outcome.

We also suggest developing comprehensive, mutually exclusive outcome codes as well as emphasizing the importance of collecting this information in enumerator training and manuals. These measures would facilitate future research on this topic.

7. References

Bates, N. (2003) "Contact Histories in Personal Visit Surveys: The Survey of Income and Program Participation (SIPP) Methods Panel", Paper presented at the annual

conference of the American Association for Public Opinion Research.

Burt, G. and Mangaroo, R. (2003) "Nonresponse Followup (NRFU) Enumerator Training", Census 2000 Evaluation H.7.

Groves, R. and Couper, M. (1998) *Nonresponse in Household Interview Surveys*, New York: John Wiley & Sons.

Moul, D. (2002) "Nonresponse Followup for Census 2000", Census 2000 Evaluation H.5.