

THE EFFECT OF QUESTIONNAIRE LENGTH ON RESPONSE

James B. Treat*

U.S. Bureau of the Census, DSSD, Room 1657-3, Washington, DC 20233-0001

KEY WORDS: Mail Rate, Coverage, Data Quality

1. Introduction

The United States Bureau of the Census conducted a test in 1995 in three sites. The test sites are Oakland, California; Paterson, New Jersey; and six parishes in Northwest Louisiana. The Census Bureau tested three long and a short form questionnaires as part of the continuing research into the effect of form length on response. Specifically, the objective is to determine the effect of variable questionnaire length has on mail response and data quality. Questionnaire length is defined as the number of questions. The mail response rate analysis is performed and compared across four questionnaire types. If the difference in the mail response rates is statistically significant across the questionnaire types, then this would result in an increase in the number of housing units in nonresponse followup for the questionnaire with the lower mail response rate. The data quality analysis examines the percent of mail returns where all the data are missing for at least one person on the questionnaire. This is defined as "complete missing person data." The data quality analysis is also performed and compared across the four questionnaire types. If the difference in the percent of questionnaires with complete missing person data is statistically significant across the questionnaire types, then this would result in an increase in the number of questionnaires (the workload) requiring either edit followup or imputation for the questionnaires with higher percent of complete missing person data.

2. Background

There are three long form questionnaires in addition to the short form questionnaire. All four questionnaires contain the seven 100 percent questions; i.e., sex, age, relationship, Hispanic origin, race, tenure (renter/owner status) and respondent's telephone number. The short form questionnaire consists of eight pages. The three long form questionnaires contain substantially different numbers of questions. The comprehensive long form (COM) contains a total of 54 questions and is 28 pages in length. This form contains almost the same content as the 1990 Decennial Census long form questionnaire. The medium long form (MED) contains a subset of the questions on the COM; 38 questions. However, the MED has the same number of pages as the COM. Finally, the abbreviated long form (ABB) contains an even smaller subset of the questions on the COM; 17 questions and is 20 pages in length. The configuration of questions on the three long forms reflects

the effort to develop differing form lengths. The combination of questions on each long form does not indicate a hierarchy of legislative requirements for the data. However, the most sensitive questions (income and citizenship) are included on all three long form questionnaires. In addition, the three long form designs do not reflect the optimum configuration for the Census 2000.

The respondent-friendly design is utilized in the development of all four questionnaires. All population questions, 100 percent and sample, for each person, are grouped together in one area. In addition, colored graphics are used along with white background areas to identify where questions are to be answered. On the 1990 long form, questions are in the following order: the 100 percent population questions for seven persons, the 100 percent and sample housing unit questions, and the sample person questions for Person 1, Person 2, Person 3, ... and Person 7. On the 1995 long forms, the questions are in the following order: the 100 percent and sample person questions for Person 1, Person 2, Person 3, ... and Person 7. The 100 percent and sample housing unit questions are included with the person questions for Person 1. In addition, the question obtaining the respondent's telephone number is included with Person 1.

3. Methodology

3.1 Mail Response Rate

The mail response rate is defined as the number of housing units which return their questionnaire prior to the start of nonresponse followup operations divided by the number of housing units in the delivery operation. Mail response is analyzed by examining individual questionnaire mail response rates and comparing the rates by form type: short, ABB, MED and COM. Treat (1995) contains a detailed description of the mail response rates.

3.2 Data Quality

Data quality is measured by the effect on item nonresponse of including the housing unit questions with Person 1 and combining the 100 percent and sample person questions for each person. Analysis of item nonresponse by question is not possible since the data from the housing unit and sample person questions were not keyed. Therefore, an alternative method is developed for analyzing the effect on data quality. The effect on data quality (missing person data) is measured by comparing the number of persons listed on the questionnaire roster (roster count) to the

number of persons with data (inside count).

In order to determine the effect of the placement of the housing unit questions, we calculated the percent of mail returns where the roster count is greater than one and the inside count is equal to one. If the percent differed by form type, then this would indicate an increase in the number of questionnaires (the workload) with missing person data, thus requiring either edit followup or imputation.

In order to determine the effect of combining the 100 percent and sample person questions, we calculated the percent of mail returns where the roster count is greater than the inside count. If the percent differed by form type, then this would indicate an increase in the number of questionnaires (workload) with missing person data, thus requiring either edit followup or imputation.

3.3 Statistical Inference

Standard errors for the form type estimates are computed using the simple random sampling jackknife variance procedure (Wolter, 1985). The analysis entailed a series of pairwise comparisons. Due to the various hypotheses being tested, all possible pairwise comparisons between the four form types (six total) are analyzed. The more comparisons that are made, the greater the potential that some of these comparisons will be declared significant when they are not. In this case, additional statistical measures are employed to control the overall error of the decision process.

The mail response and data quality analysis is carried out so that statements about the entire "family" of six comparisons are made while maintaining the 90 percent (which is a Census Bureau standard) confidence level simultaneously for all comparisons. While this procedure requires that larger differences exist before two forms are declared significant, we are at least 90 percent confident that all such decisions are correct, simultaneously. All 90 percent confidence intervals are adjusted using the procedure in Hochberg and Tamhane (1987) for comparing pairwise contrasts of the form estimates.

4. Limitations

The results are based on the test census environment and not the Decennial Census environment. During a Decennial Census, there is a national awareness of the census through national and local advertising and the media. The combination of the national and local promotion/awareness is called the Decennial Census environment and contributes to improving the response rates. The test census operated in the absence of this Decennial Census environment. Therefore, the mail response rate results do not predict levels of response in the Census 2000, rather they reflect patterns of response. In addition, the difference in the mail response rate between

the short and long form questionnaires in the 1990 Decennial Census is much smaller than what is observed in the test.

5. Results

5.1 Mail Response Rates

Table 1 contains the mail response rates by test site and form type. The mail response rate for all forms in Oakland and Paterson (the two urban sites) is 58.7 percent and 43.0 percent, respectively. The total response rate for the rural Louisiana test site is 57.1 percent.

**Table 1:
Mail Response Rates by Form Type and Test Site**

Form Type	Test Site		
	Oakland	Paterson	Louisiana
Short	60.1 (0.1)	44.4 (0.2)	58.1 (0.2)
ABB	53.7 (0.5)	37.7 (0.9)	54.2 (0.9)
MED	50.3 (0.5)	35.0 (0.9)	51.1 (0.9)
COM	48.9 (0.5)	33.9 (0.9)	50.5 (0.9)
Total	58.7	43.0	57.1

Standard Errors are in ()

Table 2 contains the comparisons which indicate the effect questionnaire length has on mail response. Table 2 contains the six possible comparisons of the mail response rate for each test site. The first three comparisons are between the short form questionnaire and each long form questionnaire. When we increase questionnaire length from the short form with seven questions to the ABB with 17 questions, there is a decrease in the mail response rate of 6.4 percentage points in Oakland, 6.8 percentage points in Paterson, and 4.0 percentage points in Louisiana. In addition, this comparison represents an increase in the questionnaire size of 12 pages. When we increase questionnaire length by 31 questions and questionnaire size by 20 pages to create the MED from the short form, there is a decrease in the mail response rate of 9.9 percentage points, 9.5 percentage points, and 7.1 percentage points in Oakland, Paterson, and Louisiana, respectively. Finally, when comparing the short form with the COM, the addition

of 47 questions and 20 pages decreases the mail response rate by 11.3 percentage points in Oakland, 10.5 percentage points in Paterson, and 7.7 percentage points in Louisiana. The reduction in the mail response rates are similar in both urban test sites. Response to the three long form questionnaires seemed to be slightly higher in the Louisiana test site. For all three test sites, the observed differences between the short form and the three long forms are statistically significant at the 90 percent confidence level.

Table 2: Comparison of Mail Response Rates by Form Type and Test Site

Form Type Comparison	Response Rate Differences		
	Oakland	Paterson	Louisiana
ABB-Short	-6.4 * (0.6)	-6.8 * (1.0)	-4.0 * (1.0)
MED-Short	-9.9 * (0.6)	-9.5 * (1.0)	-7.1 * (1.0)
COM-Short	-11.3 * (0.6)	-10.5 * (1.0)	-7.7 * (1.0)
ABB-MED	-3.4 * (0.8)	-2.7 (1.3)	-3.1 * (1.3)
ABB-COM	-4.8 * (0.8)	-3.7 * (1.3)	-3.7 * (1.3)
MED-COM	-1.4 (0.8)	-1.0 (1.3)	-0.6 (1.3)

* indicates that the difference is statistically significant at the Family Wise Error $\alpha=0.1$
Standard Errors are in ()

The next two comparisons are between the ABB and the two other long forms. The comparison between the ABB and the MED represents an increase in questionnaire length of 21 questions and questionnaire size of eight pages. In Oakland, Paterson, and Louisiana, the mail response rates decreases by 3.4 percentage points, 2.7 percentage points, and 3.1 percentage points, respectively. For Oakland and Louisiana the observed differences are statistically significant at the 90 percent confidence level. There is no evidence of a difference in the mail response rates for Paterson. The comparison between the ABB and the COM represents an increase in questionnaire length of 37 questions and questionnaire size of eight pages. The mail response rates decreases by 4.8 percentage points in Oakland, 3.7 percentage points in Paterson, and 3.7 percentage points in Louisiana. The observed differences between the ABB and the COM are statistically significant

at the 90 percent confidence level.

Finally, the last comparison is between the MED and the COM. This represents an increase in questionnaire length of 16 questions and no increase in the questionnaire size of 28 pages. In Oakland, Paterson, and Louisiana, the mail response rate decreases by 1.4 percentage points, 1.0 percentage points, and 0.6 percentage points, respectively. For all three test sites, there is no evidence that the differences are statistically significant.

5.2 Data Quality

5.2.1 Placement of the Housing Unit Questions

The 1990 Decennial Census long form questionnaire design had separate sections for the 100 percent person data, the 100 percent and sample housing unit data, and the sample person data. The design of the 1995 Census Test questionnaires moves the housing unit questions to the questions for Person 1. In addition, the 1995 questionnaire design merges the 100 percent and sample person questions for each person into a single section. The concern is that the placement of the housing unit questions could cause a reduction in data quality; i.e., an increase in missing person data, thus increasing either the edit followup workload or the imputation workload. On the short form questionnaire, seven questions are asked of Person 1 while five questions are asked of Persons 2 through 7. On the COM questionnaire, 54 questions are asked of Person 1 while 30 questions are asked of Persons 2 through 7. On the MED questionnaire, 38 questions are asked of Person 1 while 23 questions are asked of Persons 2 through 7. Finally on the ABB questionnaire, 17 questions are asked of Person 1 while 12 questions are asked of Persons 2 through 7.

The respondent might assume that the same number of questions asked for Person 1 would be asked for each of the remaining persons in the housing unit. Thus, the respondent might stop completing the questionnaire with Person 1 and return the incomplete form. In order to determine the effect the placement of the housing unit questions has on data quality (missing data), we compare the number of persons listed on the questionnaire roster (roster count) to the number of persons with data (inside count). Housing units where either count is missing and/or the roster count is zero are excluded from the analysis.

Table 3 contains the percent of mail returns by questionnaire type where the roster count is greater than one and the inside count is equal to one. The percents in Table 3 indicate when the respondent stopped completing the questionnaire with Person 1, thus analyzing the effect of the placement of the housing unit questions.

For the Oakland test site, 0.4 percent of the mail returned short forms have a roster count greater than one and an inside count equal to one. For the three long forms, less than one percent of their mail returns have a roster

count greater than one and an inside count equal to one. There are six possible comparisons between the four questionnaires. One comparison is statistically significant at the 90 percent confidence level. The difference between the ABB (0.3 percent) and the COM (0.6 percent) is -0.4 percentage points. The COM did better than the ABB at obtaining person data (partial or complete). However, note that the comparison is done with nine ABBs and 20 COMs. None of the remaining comparisons are statistically significant at the 90 percent confidence level.

In the Paterson test site, less than one percent of the mail returns have a roster count greater than one and an inside count equal to one, for all four types of questionnaires. None of the six comparisons are statistically significant at the 90 percent confidence level.

For the Louisiana test site, less than 1.2 percent of the mail returns have a roster count greater one and an inside count equal to one, for all four questionnaire types. None of the six comparisons are statistically significant at the 90 percent confidence level.

Table 3: Roster Count Greater than 1 and Inside Count Equal to 1 for Mail Return Questionnaires

Form Type	Test Site		
	Oakland	Paterson	Louisiana
Short	0.4 (0.1)	0.5 (0.1)	0.7 (0.1)
ABB	0.3 (0.1)	0.9 (0.3)	1.1 (0.6)
MED	0.4 (0.1)	0.7 (0.3)	0.5 (0.4)
COM	0.6 (0.1)	0.7 (0.3)	0.5 (0.4)

Standard Errors are in ()

5.2.2 Combining 100 Percent and Sample Person Questions

The 1990 Decennial Census long form questionnaire first asked the 100 percent person questions for all seven persons. After the 100 percent person questions are answered, there is a separate section for each person to supply the data for the sample person questions. The assumption is that grouping the questions in this way would maximize the number of forms we receive back with complete 100 percent person data.

The design of the 1995 Census Test long form

questionnaires combine the 100 percent person questions with the sample person questions. For each person, the respondent would answer all the person questions, 100 percent and sample. With the design of the 1995 Census Test long forms, the respondent might stop completing the questionnaire at any point in the questionnaire. The concern is that combining the 100 percent and sample person questions for each person, in tandem with the placement of the housing unit questions with Person 1, might result in a decrease in the number of persons for which we obtain any data (increase in complete missing person data), especially the 100 percent person questions.

In order to determine the effect of combining the 100 percent and sample person questions we compare the number of persons on the roster (roster count) to the number of persons with data (inside count) for the mail return questionnaires. When the roster count is greater than the inside count, this indicates that there are housing units with complete missing person data, 100 percent and sample person data. One possible reason for the complete missing person data could be the placement of the sample person questions with the 100 percent person questions. Note that housing units where either count is missing and/or the roster count is zero are excluded from the analysis.

Table 4 contains the percent of mail returns when the roster count is greater than the inside count by form type and test site. Results for the short form questionnaire are presented as a benchmark when comparing to the three long form questionnaires.

Table 4: Roster Count Greater Than the Inside Count for Mail Return Questionnaires

Form Type	Test Site		
	Oakland	Paterson	Louisiana
Short	3.0 (0.1)	3.4 (0.2)	4.1 (0.3)
ABB	0.8 (0.2)	1.9 (0.5)	2.0 (0.7)
MED	1.2 (0.2)	2.1 (0.5)	1.1 (0.5)
COM	1.4 (0.2)	2.0 (0.5)	1.8 (0.7)

Standard Errors are in ()

In the Oakland test site, the roster count is greater than the inside count for 3.0 percent of the short forms, 0.8 percent of the ABBs, 1.2 percent of the MEDs, and 1.4

percent of the COMs.

In the Paterson test site, the roster count is greater than the inside count for 3.4 percent of the short forms, 1.9 percent of the ABB, 2.1 percent of the MEDs, and 2.0 percent of the COMs.

Finally in the Louisiana test site, the roster count is greater than the inside count for 4.1 percent of the short forms, 2.0 percent of the ABBs, 1.1 percent of the MEDs, and 1.8 percent of the COMs.

The concern is with the possible reduction in content resulting from an increase in missing person data. In order to determine the impact, we compare the percent of mail return questionnaires where the roster count is greater than the inside count across form type. There are six possible comparisons between the four form types. Table 5 contains the comparisons.

Table 5: Comparing the Percent of Mail Return Questionnaires where the Roster Count is Greater than the Inside Count by Form Type and Test Site for Mail Return Questionnaires

Form Type Comparison	Test Site		
	Oakland	Paterson	Louisiana
ABB-Short	-2.2 * (0.2)	-1.6 * (0.5)	-2.2 * (0.8)
MED- Short	-1.7 * (0.2)	-1.3 * (0.6)	-3.1 * (0.6)
COM-Short	-1.5 * (0.2)	-1.4 * (0.6)	-2.4 * (0.7)
MED-ABB	0.4 (0.2)	0.2 (0.7)	-0.9 (0.9)
COM-ABB	0.6 * (0.3)	0.2 (0.7)	-0.2 (1.0)
COM-MED	0.2 (0.3)	-0.1 (0.8)	0.7 (0.8)

* indicates that the difference is statistically significant at the Family Wise Error $\alpha=0.1$
Standard Errors are in ()

The first three comparisons are between the short form and three long form questionnaires. In all three test sites, the three long form questionnaires have a statistically significant lower percentage of questionnaires with the roster count greater than the inside count. The differences indicate that the short form has a larger percentage of forms with complete missing person data than any of the three long forms.

The next three comparisons are among the three long form questionnaires. In the Oakland test site, the difference of 0.6 percent between the COM and ABB is statistically significant at the 90 percent confidence level. This difference indicates that the COM has a larger percentage of forms with complete missing person data than the ABB. None of the remaining comparisons are statistically significant at the 90 percent confidence level.

6. Conclusions

We observed that as questionnaire length increases, total mail response rate decreases. Between the short form questionnaire and all of the long form questionnaires, we observed a statistically significant decrease in the mail response rate. However, the decreases in mail response rate seem to level off as we compare long form mail response rates. This is illustrated with two separate comparisons in Oakland. The first comparison is between the ABB and the short form. We observed a statistically significant 6.4 percentage point decrease in the mail response rate. This comparison represents the addition of ten questions to a seven question form (short form) and the addition of twelve pages to an eight page form. The next comparison is between the MED and COM. We observed a 1.4 percentage point decrease in the mail response rate which is not statistically significant. This comparison represents an increase in questionnaire length of 17 questions from a 38 question form (MED) to a 54 question form (COM) and no additional pages. Therefore, the reduction in the mail response rate seems to be dependent on the number of questions added to the form, the overall number of questions on the form and the number of pages. In Paterson and Louisiana we observed similar results. It is important to remember that during a decennial census, the difference between the short and long form questionnaires is smaller than what we observed in the 1995 Census Test.

The analysis of the placement of the housing unit questions indicates that, for the majority of the comparisons, there is no evidence of a significant difference. The only comparison which shows a statistically significant difference is in Oakland between the COM and the ABB. This comparison is done using 20 COMs and nine ABBs. None of the remaining comparisons in the three test sites are statistically significant at the 90 percent confidence level. Therefore, the placement of the housing unit questions with Person 1 appears not to affect the amount of missing person data differently among the four questionnaire types.

The analysis of combining the 100 percent and sample person questions also supports the design. The comparison between the short form and the COM shows that the COM did a better job at obtaining person data (partial or complete) than the short form. This is the reverse of what we expected.

7. Recommendations

Based on the results, the Census Bureau should continue using the COM questionnaire. We observed in the 1995 Census Test only 4.8, 3.7 and 3.7 percentage point decreases between the ABB and the COM in Oakland, CA, Paterson, NJ and Louisiana, respectively. During the 1990 Decennial Census, the difference in response rate between the short and long form questionnaires is only 4.5 percentage points (Treat, 1993). Given the census environment in 2000, we might not observe a difference between an ABB and a COM.

Based on these results, the placement of the housing unit questions with Person 1 and combining the 100 percent and sample person questions does not affect the obtainment of person data across the forms. Therefore, the Census Bureau could continue using this design without the loss of person data and without an increase in the number of questionnaires requiring either edit followup or imputation.

8. References

Hochberg, Y. and A.C. Tamhane. 1987. *Multiple Comparison Procedure*. New York, John Wiley and Sons.

Treat, James B. 1993. "1993 National Census Test Appeals and Long Form Experiment (ALFE). Long Form Component: Final Report." *DSSD 2000 Census Memorandum Series #E-63*. U.S. Bureau of the Census, November 17.

Treat, James B. 1995. "Definition of Mail Check-in, Initial Questionnaire, Response, Completion and Estimated Return Rates for the 1995 Census Test." *DSSD 1995 Census Test Memorandum Series #L-2 Revision 1*. U.S. Bureau of the Census, July 23.

Wolter, Kirk. 1985. *Introduction to Variance Estimation*. New York, Springer-Verlag.

* This paper reports the general results of research undertaken by Census Bureau staff. The views expressed are attributable to the author and do not necessarily reflect those of the Census Bureau.