

The Effect of Interview Length on Data Quality in the Consumer Expenditure Interview Survey October 2014

Brian T. Nix¹

U.S. Bureau of Labor Statistics, Office of Prices and Living Conditions
2 Massachusetts Avenue NE, Washington, DC 20212

Nix.Brian@bls.gov

Abstract

The Consumer Expenditure Interview Survey is an ongoing panel survey of U.S. households in which sample households receive the same survey in a series of four interviews, each typically lasting about an hour. Survey length is one element thought to negatively affect survey results, by increasing respondents' perception of burden. Because of the high burden associated with the survey, and the steady decrease in the survey's response rates, measures are being taken to shorten the interviews. This paper examines the effect of the interview's length on response rates, reported expenditure values, and other measures of data quality.

Key Words: burden reduction, Consumer Expenditure Survey, data quality, response rate

1. Introduction

The Consumer Expenditure (CE) Survey is a nationwide household survey conducted by the U.S. Bureau of Labor Statistics (BLS) to find out how Americans spend their money. The CE survey consists of two sub-surveys, an Interview survey and a Diary survey, collected for the BLS by the U.S. Census Bureau. The purpose of the Diary survey is to collect detailed expenditure data on small, frequently purchased items such as food and apparel. The purpose of the Interview survey is to collect detailed expenditure data on large items such as property, automobiles, and major appliances, as well as on recurring expenses such as rent, utility bills, and insurance premiums. The data from the two surveys are then combined to provide a complete picture of consumer expenditures in the United States.

This paper looks at only the Interview survey in an effort to determine what effect the amount of time a respondent spends on an interview has on reported expenditures and response rates in subsequent interviews. By looking at the amount of time respondents spend on an interview in one quarter, and then looking at their response rates and reported expenditures in the next quarter, it was found that the length of an interview does not affect response rates, but it does affect the amount of expenditures reported. Longer interviews tend to have more expenditures reported, but shorter interviews lead to more expenditures being reported in the following interviews.

2. Data Description

Each household in the Interview survey is interviewed four times, once per quarter, over a one-year period.² Data are collected from approximately 7,000 households per quarter, for a total of approximately 28,000 quarterly interviews per year. The data used in this paper

¹ Any opinions expressed in this paper are those of the author and do not constitute policy of the Bureau of Labor Statistics.

² There is also an additional, initial interview, during which demographic information is collected, as well as some expenditure information, but the data from this interview are only used internally for bounding purposes – to classify the unit for analysis, as well as to prevent duplicate reporting. Thus, the interviews in our database are numbered 2 through 5.

come from the Interview survey over the eight-year period 2005-2012, which means over 220,000 interviews were used.

As each interview in the survey represents one quarter, or three months, the total reported expenditures can be multiplied by four to give us annualized expenditures. These annualized expenditures are estimates of the total annual expenditures of the household. **Graph 1** shows the annualized expenditures for the 220,000+ interviews completed in 2005-2012 to have a log-normal distribution. This is to be expected since log-normal distributions are frequently found in economic data such as the income of a population. Over these eight years, the annual expenditures reported in the Interview Survey had a mean of \$46,639; a median of \$35,545; and a mode of \$23,000. When separated by interview number the data are similar, as shown in the following table:

Annual Expenditures by Interview Number

Interview Number	Mean	Median	Mode ³
2	\$46,870	\$35,734	\$23,000
3	\$46,463	\$35,461	\$24,000
4	\$46,351	\$35,368	\$22,000
5	\$46,861	\$35,619	\$26,000
All (2-5)	\$46,639	\$35,545	\$23,000

The amount of time respondents spend on interviews also follows a log-normal distribution (**Graph 2**). The distribution of interview lengths has a mean of 59 minutes; a median of 53 minutes; and a mode of 46 minutes. The distribution is similar when the data are examined by interview number. The four interviews had average times ranging from 54 to 65 minutes, with smaller medians and modes. The data are summarized in the following table:

Interview Length by Interview Number

Interview Number	Mean	Median	Mode
2	65	58	52
3	55	50	39
4	54	49	34
5	63	57	43
All (2-5)	59 minutes	53 minutes	46 minutes

Interviews 2 and 5 are longer, on average, because they contain additional questions about income that the other two interviews do not.

3. Reported Expenditures as a Function of Interview Length

In order to analyze the effectiveness of our survey, paradata – administrative data about the survey – have been collected as a part of the Interview survey since 2003. These paradata include the number of contact attempts, the length of the interview, and the primary method of contacting the respondent. This paper focuses on the number of contact attempts and the length of the interview.

³ After rounding all total expenditures to the nearest thousand dollars.

The length of the interview is automatically recorded on the field representative's laptop computer, with no special effort required. Timing starts when he or she opens the software, and it ends when the software is closed. The number of contact attempts comes from the survey's "Contact History Instrument," which is a computer application in which field representatives record information about every attempt they make to contact respondents. The number of contact attempts comes from counting the number of attempts recorded in the instrument.

There is a strong positive correlation between the amount of time spent on an interview and the total dollars of expenditures reported. **Graph 3** shows both the mean and median annualized expenditures, as a function of interview length. As expected, the mean of the expenditures is greater than the median, indicating another log-normal distribution. As the graph indicates, the expenditures increase rapidly at first, but then as the interview length becomes longer the expenditures increase at a slower rate.

Although the two variables are correlated, there is still a question of which is the cause and which is the effect. Spending more time with respondents could cause them to recall more expenditures, or respondents with more expenditures could cause the interview to last longer. The next several graphs attempt to determine which is the cause and which is the effect.

Graph 4 shows the relationship between interview length in one interview and the expenditures reported in the *next* interview. The graph is based on all households that successfully completed two consecutive quarterly interviews. It shows that the amount of expenditures reported is an increasing function of interview length. As interview length increases in one quarter, reported expenditures increase for the same household in the next quarter. Again, as interview length increases, the expenditures increase at a decreasing rate.

Graph 5 shows the *change* in the average annualized expenditures from one interview to the next based on interview length. Again, the graph is based on all households that successfully completed two successive quarterly interviews. It shows that the change in expenditures is a decreasing function of interview length. Those households who experienced shorter interviews (under 60 minutes) in one quarter tended to report higher expenditures in the next interview, while households who experienced longer interviews (over 60 minutes) in one quarter tended to report lower expenditures in the next interview. For example, when an interview only lasted 15 minutes, the graph shows the expenditures reported in the next interview were about \$3,000 higher on average; however, when an interview lasted 120 minutes, the graph shows the expenditures reported in the next interview were on average about \$3,000 lower.

Graph 6 shows the same thing, but where the change is expressed in relative terms instead of absolute terms (in percentages instead of dollars). These graphs show a negative correlation between the length of an interview and the change in the expenditures reported in the next interview. When interviews are less than one hour long the next interview tends to have an increase in expenditures; and when interviews are more than one hour long the next interview tends to have a decrease in expenditures. For example, when an interview only lasted 15 minutes, the graph shows the expenditures reported in the next interview were about 12% (\$3,000) higher on average; however, when an interview lasted 120 minutes, the graph shows the expenditures reported in the next interview were on average about 4% (\$3,000) lower.

In all four of these graphs, the variance increases noticeably for interview lengths over 90 minutes. This is a small sample size issue caused by the fact that only a small number of respondents have interview lengths over 90 minutes. **Graph 2** shows that only about 15% of all respondents have interview lengths over 90 minutes, and the mean interview length was only 59 minutes.

4. Response Rate as a Function of Interview Length

Another potential effect of interview length is the response rates of the next interview. **Graph 7** shows interview length has very little effect on the response rates of subsequent interviews. A successful interview (in the 2nd through 4th interview) is followed by another successful interview (in the 3rd through 5th interview) approximately 93% of the time, and that rate is the same regardless of the length of the interview. The response rates differ a little by interview number, but the general principle still holds that the length of an interview has very little effect on response rates in subsequent interviews.

5. Contact Attempts as a Function of Interview Length

Another measure of data quality is the number of contact attempts needed to obtain a completed interview. The number of contact attempts was shown to be related to data quality by Safir and Tan (2009), who found that reporting quality in the CE Interview Survey decreases consistently with an increase in number of contact attempts. The number of contact attempts has been recorded for our survey continuously since 2006, so there are seven years of data available through 2012. **Graph 8** shows that the average number of contact attempts increases with the length of the interview. This correlation shows that shorter interview lengths are related to the need for fewer contact attempts, which makes the survey more efficient, in addition to increasing the quality of the data for the survey.

In order to see the effect of interview length, it is necessary to compare the difference from one interview to the next. **Graph 9** shows the average number of contact attempts in the next interview as a function of current interview length. **Graph 10** shows the average reduction in number of contact attempts necessary for the same household from the current interview to the next interview, based on current interview length. Together, these graphs show that, for interviews lasting at least one hour, longer interview length is related to fewer contact attempts necessary for the next interview.

6. Conclusions

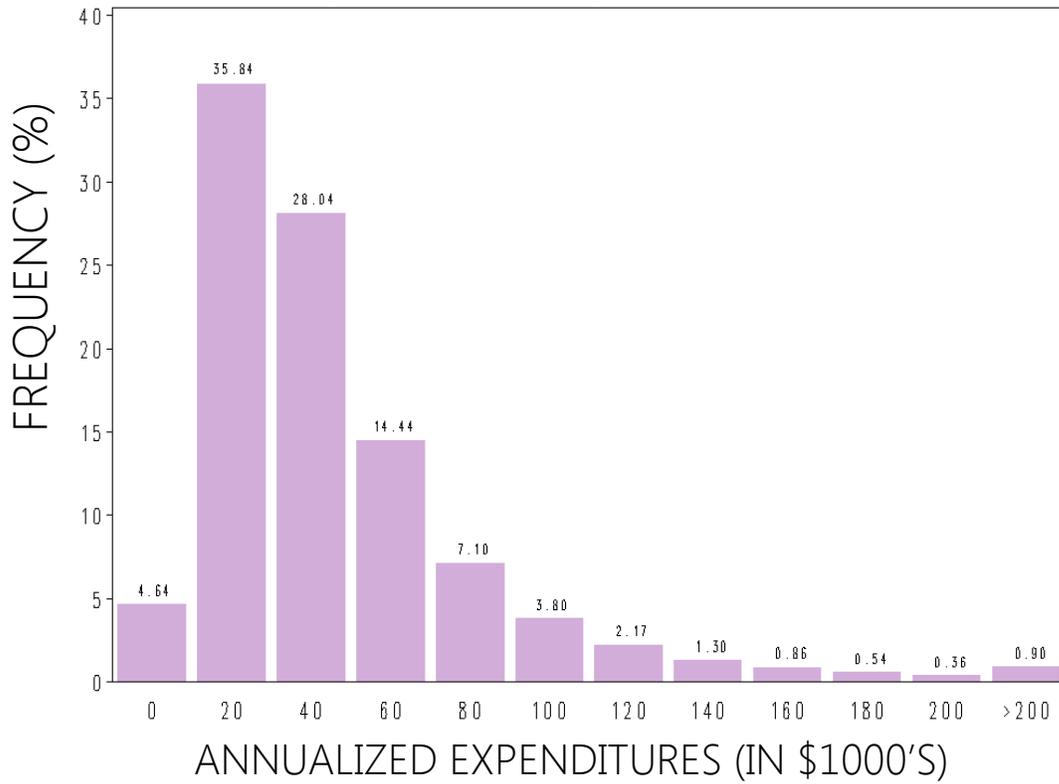
This paper has looked at the effect of the length of the interview survey on its response rates, reported expenditures, and number of contact attempts. By looking at the length of each interview and comparing the response rates, the number of contact attempts, and the reported expenditures in those interviews with the same results in the next interview, it was found that interview length does affect the amount of expenditures reported, but it does not affect the response rate. Longer interviews have more expenditures reported in the current interview, but less expenditures in the next interview.

However, longer interviews do not appear to have a negative impact on response rate or number of contact attempts in the next interview. These results show that longer interviews do not adversely affect the quality of the data collection process, as much as they affect the quality of the data collected. The CE survey is currently undergoing a redesign, and the information discovered in this paper should be helpful in that effort.

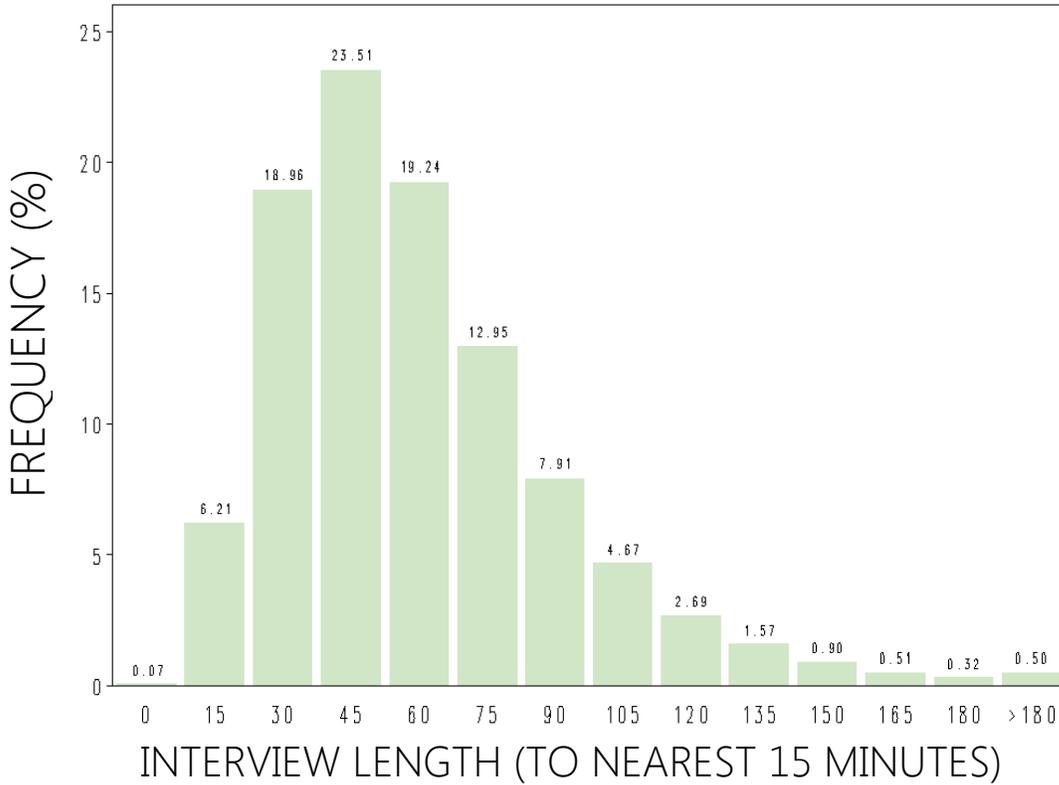
References

Safir, A., and Tan, L. 2009. "Using Contact Attempt History Data to Determine the Optimal Number of Contact Attempts." Paper presented at the Annual Meeting of the American Association for Public Opinion Research, Hollywood, FL.

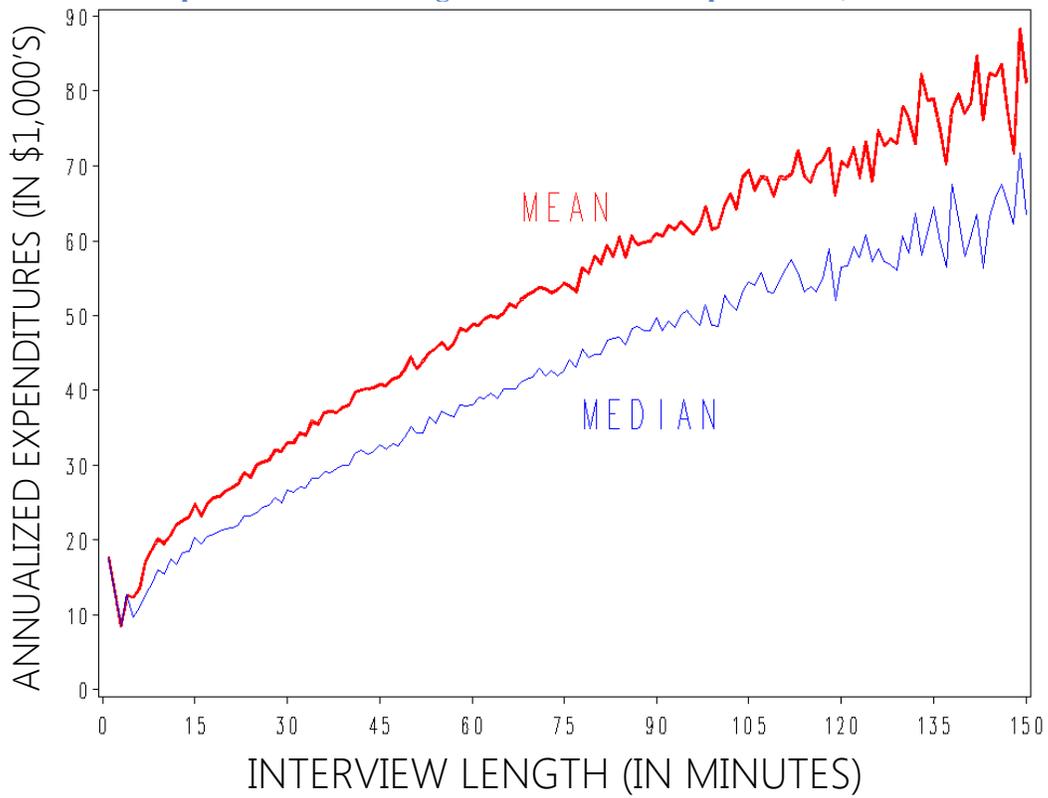
Graph 1 – Annualized Expenditures, 2005-2012



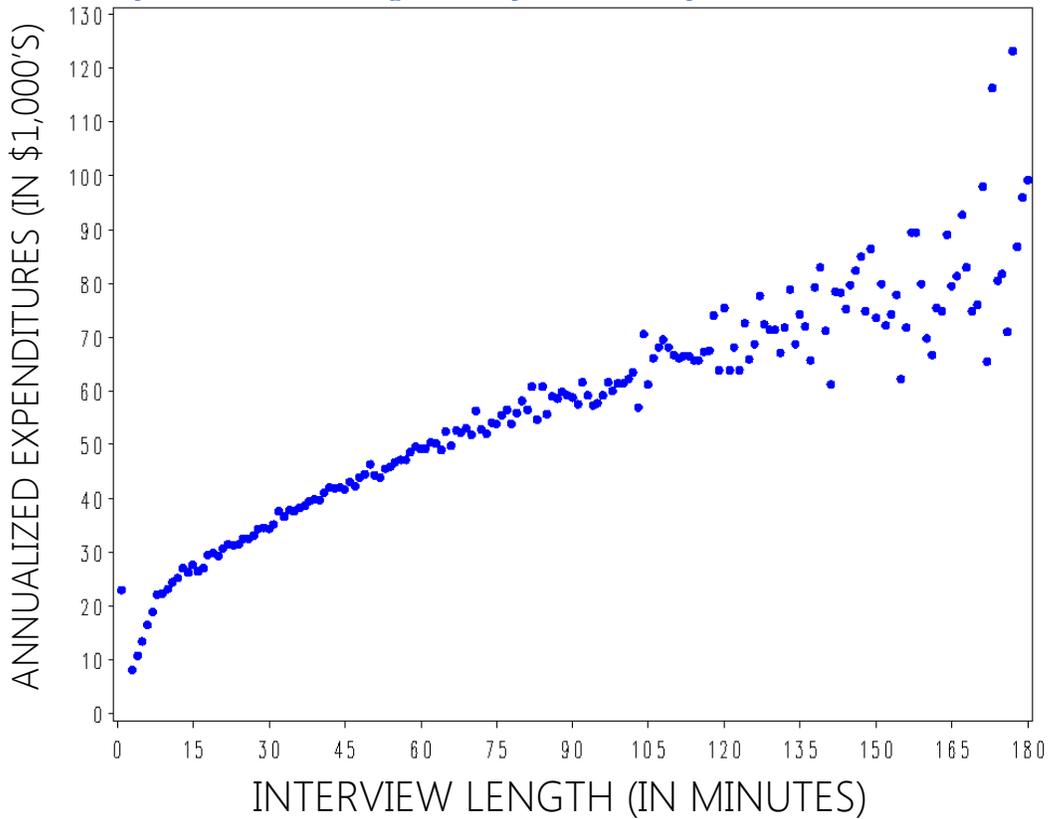
Graph 2 – Interview Length, 2005-2012



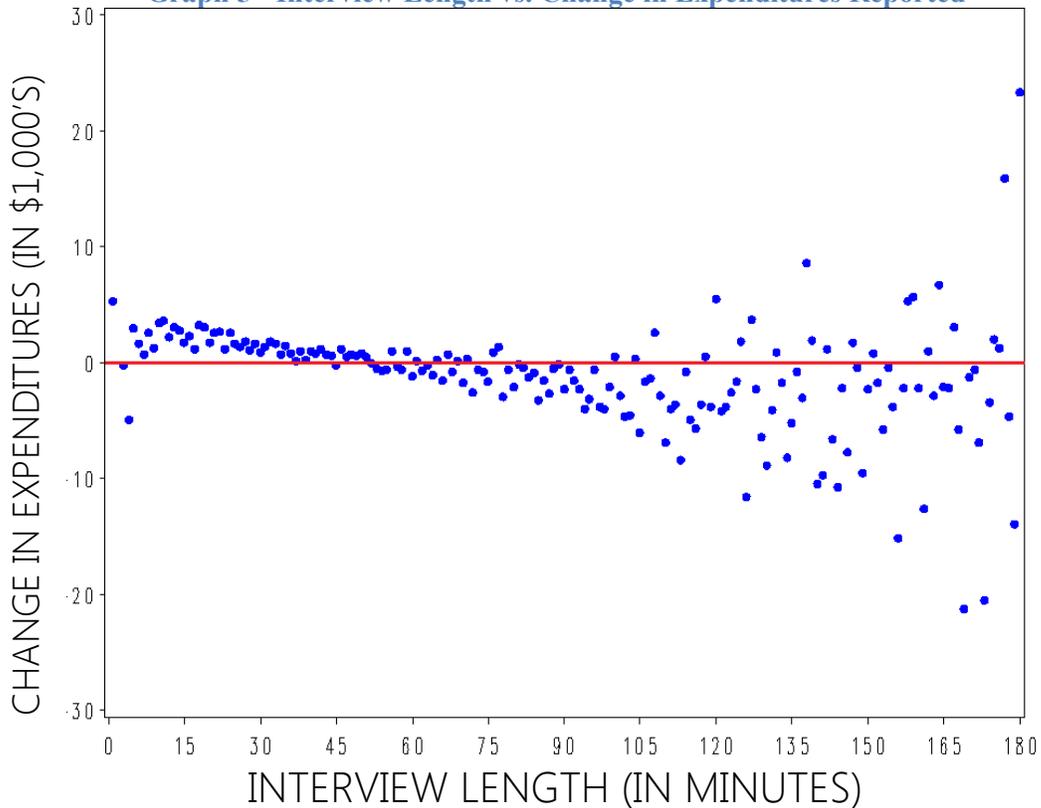
Graph 3 - Interview Length vs. Annualized Expenditures, 2005-2012



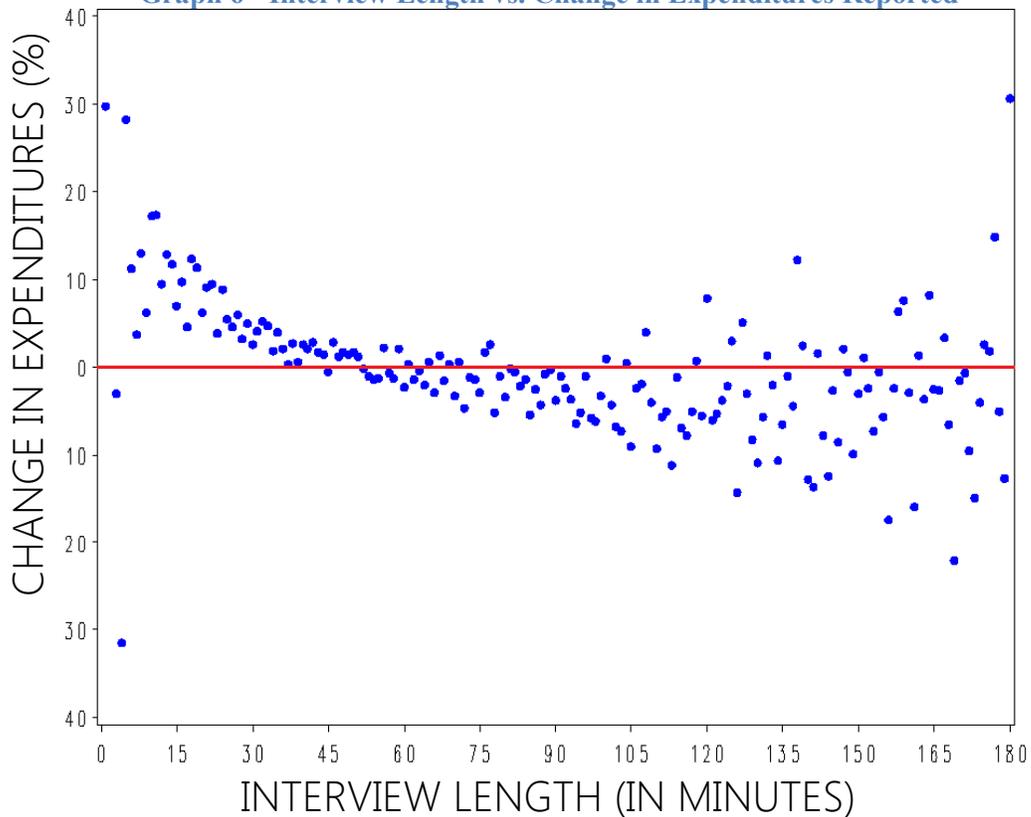
Graph 4 - Interview Length vs. Expenditures Reported in the Next Interview



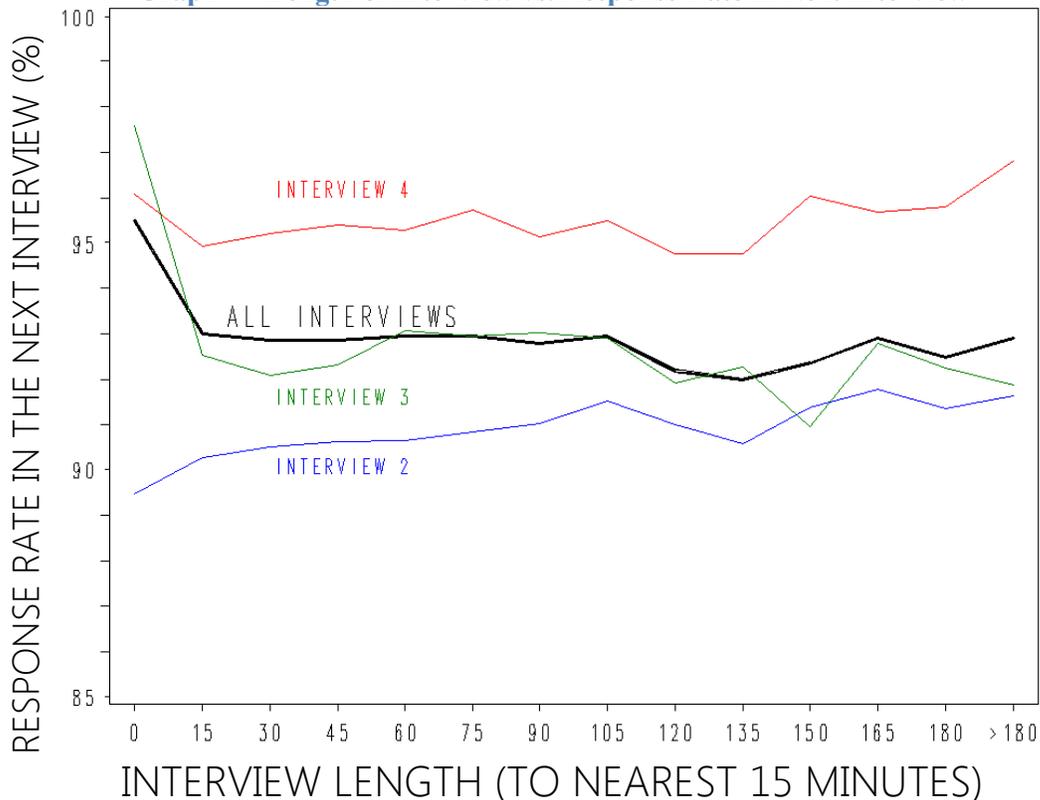
Graph 5 - Interview Length vs. Change in Expenditures Reported



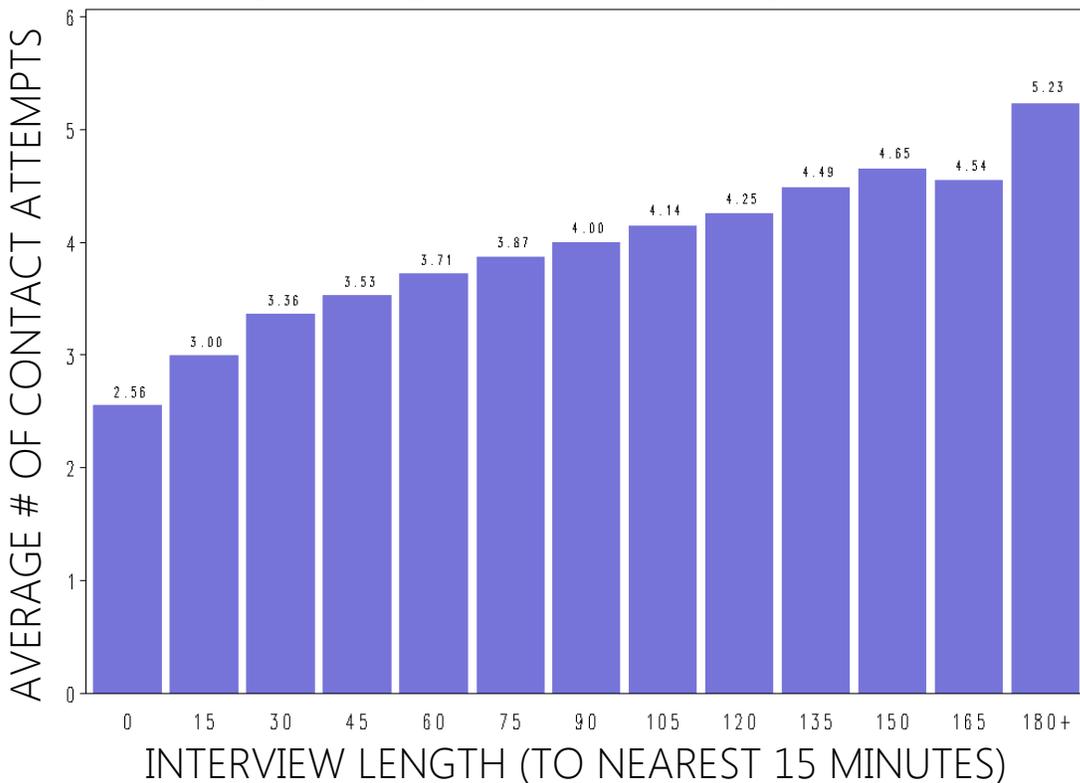
Graph 6 - Interview Length vs. Change in Expenditures Reported



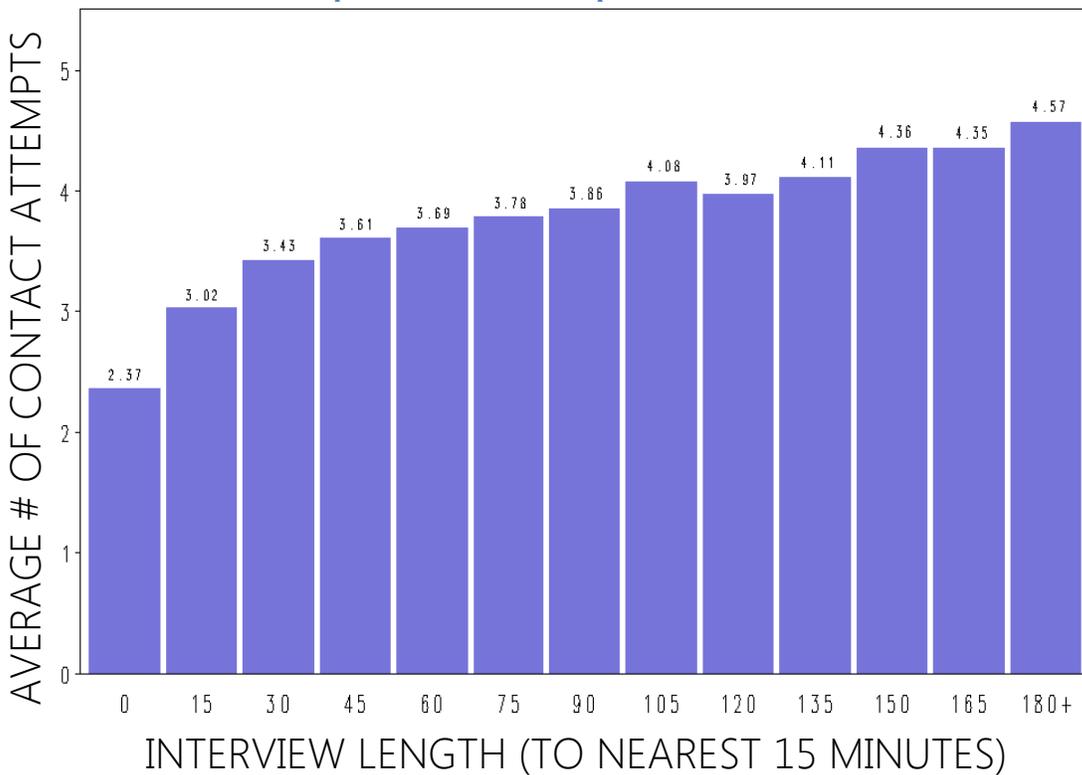
Graph 7 - Length of Interview vs. Response Rate in Next Interview



Graph 8 - Average Number of Contact Attempts, 2006-2012



Graph 9 - Contact Attempts for Next Interview



Graph 10 - Average Change in Number of Contact Attempts

