Nonresponse Bias in the Consumer Expenditure Survey: A Case Study

Susan L. King[1], Jennifer Edgar[2], Jeffrey M. Gonzalez[2], Boriana Chopova[1], and Lucilla Tan[1]

The Consumer Expenditure Interview Survey (CE) is a nationwide survey conducted by the U.S. Bureau of Labor Statistics to estimate the expenditures made by American households. The response rate for the survey has varied between 74.5 and 78.6 percent from 2002 to 2007. Response rates in this range indicate a risk for nonresponse bias. This paper is a synthesis of four studies undertaken to study whether nonresponse introduces bias into the survey estimates. The four studies are: (1) A comparison of response rates between subgroups of the survey's sample, (2) A comparison of respondent demographic characteristics between the CE and the American Community Survey, (3) An analysis of nonresponse bias using 'harder-to-contact' respondents as proxies for nonrespondents, and (4) An analysis of nonresponse bias using intermittent respondents and attritors as proxies for nonrespondents. Collectively, the studies show no meaningful bias in the survey's estimates even though the nonresponse is not missing completely at random.

Key words: Continuum of resistance, intermittent respondent, missing data, panel survey, proxy nonrespondent

1. Introduction

The purpose of this paper is to present a case study that investigated nonresponse bias in the Consumer Expenditure Interview Survey (CE). The CE is a nationwide household survey conducted by the U.S. Bureau of Labor Statistics (BLS) to estimate the expenditures made by American households. It is a federal statistical survey with a response rate ranging from 74.5 percent to 78.6 percent between collection years 2002 and 2007. Low response rates are a concern because they indicate risk for nonresponse bias. The CE is a rotating panel survey in which approximately 15,000 households are visited each quarter of the year, and each household is contacted for an interview every three months for five consecutive quarters. Expenditure information from the first interview is not used in the CE's

[1] Consumer Expenditure Survey, U.S. Bureau of Labor Statistics, 2 Massachusetts Ave. NE, Washington DC 20212. Email: king.susan@bls.gov

[2] Office of Survey Methods Research, U.S. Bureau of Labor Statistics, 2 Massachusetts Ave. NE, Washington DC 20212. Email: edgar.jennifer@bls.gov

published estimates. Instead, the first interview is used only for inventory and 'bounding' purposes, which addresses a common problem in which survey respondents tend to report expenditures to have been made more recently than they were actually made. Only expenditure information from the second through fifth interviews is used in the published estimates.

To evaluate nonresponse bias in the CE, four studies were completed:

1. A comparison of response rates between subgroups of the survey's sample.

2. A comparison of socio-demographic characteristics to an external data source (the American Community Survey, or ACS)

3. An analysis of nonresponse bias using 'harder-to-contact' respondents as proxies for nonrespondents.

4. An analysis of nonresponse bias using intermittent respondents and attritors (survey dropouts) as proxies for nonrespondents.

These studies were designed to answer the questions: (1) Are the data in the CE missing completely at random (MCAR)? (2) What are the demographic characteristics of the nonrespondents? and (3) What is the level of nonresponse bias in the CE?

The first step in most nonresponse bias studies is to understand the missing data mechanism. Data are said to be 'missing completely at random' if the mechanism that produces the missing values is unrelated to the values of the data themselves (Little and Rubin 2002). More precisely, data are 'missing completely at random' if their pattern of 'missing-ness' is independent of the data's actual values and the values of any other variables. Another mechanism for generating missing data is 'missing at random' (MAR). Data are said to be 'missing at random' (MAR) if the probability of data missing for a particular variable is unrelated to the value of that variable after controlling for other variables in the analysis. MAR is a weaker assumption than MCAR and it is impossible to verify whether the MAR condition is satisfied. The values of the missing data are unknown and thus a systematic comparison of observations with and without the missing variable is not possible (Allison 2002).

Nonresponse bias is often associated with the data not being MCAR. One common method of determining whether the data are MCAR is to examine differences between respondents and nonrespondents on variables that are collected for both groups and that are associated with the values of the survey data. Any differences suggest that the

missing data may not be MCAR. Frame and stratification variables are frequently used to test for MCAR in this approach (e.g., Brick and Bose 2001; Dixon 2001 and 2004; Duncan and Hill 1989; Pearl and Fairley 1985; and Purdie et al. 2002). Another common approach is to compare characteristics of respondents to characteristics of an external population. Again, differences between the two groups suggest that the missing data may not be MCAR. The first two studies followed these approaches, but all four studies included MCAR analyses.

After answering the MCAR question, the next step is to estimate the amount of nonresponse bias in the survey's expenditure estimates. Since by definition no data are collected from nonrespondents, one approach is to partition the respondent sample into subgroups that exhibit different nonresponse bias characteristics (Groves 2006), and then select one of the subgroups to serve as 'proxies' for the nonrespondents. Once a set of proxy nonrespondents is selected, a dataset can be created with expenditure and socio-demographic information that is representative of the nonrespondents, and the level of nonresponse bias can be estimated. The third and fourth studies in this report use this approach. The third study uses 'harder-to-contact' respondents as proxy nonrespondents and the fourth study uses intermittent respondents and attritors as proxy nonrespondents. Both studies draw on the 'continuum of resistance' theory. Sample units are ordered by the amount of interviewer effort exerted to obtain a completed interview, and those requiring the most effort are chosen to serve as the proxy nonrespondents (Groves 2006). In both of these studies nonresponse bias was computed for total expenditures and 13 expenditure subcategories, ranging from regular monthly expenses such as housing payments to infrequent and highly variable expenses such as those for education.

2. Methodology: Common Approaches Across Studies

2.1. Data

For comparability of results, the four studies used a common data file for analysis. This common file includes 15 months of data, from April 2005 through June 2006. Data on respondent contacts used in two of the studies come from the Contact History Instrument (CHI), an interviewer completed instrument which collects data on every contact attempt made for each CE case. The CHI data includes information about the date, time, method, and outcome of each contact. The unit of analysis in these studies is the consumer unit (CU), which in most cases is a household. The common data file consists of one record per wave (or interview) per CU for Waves 1, 2, and 5.

Waves three and four were excluded for simplicity, and with the expectation that any bias found would be consistent in these waves.

In each record, there were CU-level variables as well as variables for respondent characteristics. In the event of a noninterview, the respondent characteristics were missing except for interviewer-recorded values on member race, CU size, and housing tenure. These interviewer-recorded values are often based on speaking to neighbors or from the interviewer's inference. The sample size of the common data file is shown in Table 1.

Table 1. Sample size of common data file by wave

| Wave | Completed Interview | Type A Noninterview | | | | Total Eligible Cases |
		Refused	No one Home	Temporarily Absent	Other	
1	9,795	1,640	459	166	391	12,451
2	9,432	2,017	407	138	378	12,372
5	9,547	2,225	279	100	354	12,505
Total	28,774	5,882	1,145	404	1,123	37,328

Records for other CU members were added to the common data file for member-level analyses in the second study, 'Comparison of the Characteristics of CE Respondents to External Data.'

Variables collected in the CHI were added to the common data file for analysis in the third study, 'Harder-to-Contact Respondents as Proxies for Nonrespondents.' These variables provide information about the number of times an interviewer contacted the CU prior to completing the interview.

Several studies required Wave 1 demographic data. Since Wave 1 is primarily a bounding interview and is not used in published estimates, the data are not subjected to the regular cleaning, editing, and variable creation activities that are performed on data in Waves 2 through 5. To use these data in the common data file, Wave 1 demographics were processed following the production specifications used in Waves 2 through 5. Slight modifications were necessary for race and housing tenure (owner or renter) variables. The values for race and housing tenure created for this series of studies were tested against those generated from production data for Waves 2 through 5. For each wave,

the match rates between the created code and the production code for race and tenure were over 99%, while the match rates for the size of the CU ranged from 85% to 89%.

2.2. Weighting

The CE's sample design is a nationwide probability sample of addresses. The selection of addresses for the survey begins with the definition and selection of geographic areas called "primary sampling units" (PSUs), which are small groups of contiguous counties. CE's sample in 2006 consisted of 91 PSUs randomly selected to represent the entire nation. Within each of these PSUs a systematic sample of addresses is drawn. Most addresses consist of one CU, but some addresses have more than one CU. Each interviewed CU represents itself as well as other CUs and therefore must be weighted to properly account for all CUs in the population. The U.S. Census Bureau selects the sample and provides the base weights, which are the inverse of the CU's probability of selection. A Primary Sampling Unit (PSU) is a geographic area of several connected counties. The average number of counties in a PSU is five. Each CU in a PSU has the same base weight. BLS makes three types of adjustments to the base weights: an adjustment if the field representative finds multiple housing units where only a single housing unit was expected, a noninterview adjustment; and a calibration adjustment. These weight adjustments are made to each individual CU. The noninterview adjustment accounts for nonresponse by increasing the weight of the respondents in socio-demographic classes that are associated with nonresponse. Calibration adjusts the weights to Census population controls in order to account for frame under-coverage. The noninterview and calibration adjustments are more significant than the adjustment for multiple housing units.

All of the studies use base weights. The two proxy nonrespondent studies and the study comparing response rates across subgroups only used base-weighted data, but the study comparing CE respondents to external data study used all three weights (base weights, noninterview adjustment weights, and the final calibration weights).

2.3. Relative Nonresponse Bias Equation

For the estimates of nonresponse bias in the two proxy nonrespondent studies, relative nonresponse bias was computed instead of an estimate of nonresponse bias. The reason is that the dollar amounts vary substantially across expenditure categories; thus, making comparisons among them difficult. Relative nonresponse bias is a more appropriate statistic for comparisons across categories. The following formula was used to compute relative nonresponse bias, denoted as $\mathrm{Re}\,lBias\left(\overline{Z}_R\right)$, in the base-weighted sample mean:

$$\mathrm{Re}\,lBias\left(\overline{Z}_R\right) = \frac{B\left(\overline{Z}_R\right)}{\overline{Z}_T} = \frac{\overline{Z}_R - \overline{Z}_T}{\overline{Z}_T} = \left(\frac{N_{NR}}{N}\right)\left(\frac{\overline{Z}_R - \overline{Z}_{NR}}{\overline{Z}_T}\right) \tag{1}$$

where:

$B\left(\overline{Z}_R\right)$ is the nonresponse bias in the base-weighted respondent sample mean;

$\overline{Z}_R$ is the base-weighted respondent mean of expenditures;

$\overline{Z}_T$ is the base-weighted all CU mean of expenditures;

$\overline{Z}_{NR}$ is the base-weighted proxy nonrespondent mean of expenditures;

N_{NR} is the base-weighted number of proxy nonrespondent CUs; and

N is the base-weighted total number of CUs.

2.4. Variance Estimation

CE uses Balanced Repeated Replication (BRR) to calculate variance estimates. In this method, the PSUs are divided into 43 strata and the sample units within each stratum randomly assigned to one of two half samples. The full sample variance is calculated using the calibration-adjusted weights from completed interviews (Wolter 1985). There are several issues in using BRR in this project. Any change, addition or deletion, of consumer units requires re-estimating the noninterview and calibration adjusted weights. In addition, calibration weights are not calculated for Type A noninterviews, which occur when no interview is completed at an occupied eligible housing unit. The

information collected from Type A Interviews is used to calculate the nonresponse adjusted weights for completed interviews. Therefore, in this project, estimates of means and frequencies were made using two procedures from SAS®9, designed for complex surveys, PROC SURVEYMEANS and SURVEYFREQ.

The variance for relative nonresponse bias does not have a closed-form solution, so an estimate was calculated for each expenditure category using the random groups method (Wolter 1985). With the number of random groups equal to 10, the variance formula is as follows:

$$\mathrm{var}_{rg}\left(\tilde{\theta}_j\right) = \frac{1}{10(10-1)}\sum_{k=1}^{10}\left(\hat{\theta}_{k,j} - \tilde{\theta}_j\right)^2 \tag{2}$$

where:

$\hat{\bar{Z}}_{R,k,j}$ is the respondent sample mean on expenditure category j for random group k;

$\hat{\bar{Z}}_{T,k,j}$ is the total sample mean on expenditure category j for random group k;

$\hat{\theta}_{k,j} = \dfrac{\hat{\bar{Z}}_{R,k,j} - \hat{\bar{Z}}_{T,k,j}}{\hat{\bar{Z}}_{T,k,j}}$ is the relative bias on expenditure category j for random group k; and

$\tilde{\theta}_j = \dfrac{1}{10}\sum_{k=1}^{10}\hat{\theta}_{k,j}$ is the average of the relative bias on expenditure category j over all 10 random groups.

2.5. Significance Test

A 95% confidence interval of the relative nonresponse bias of expenditures was computed as follows:

$$\hat{\theta}_j \pm t_{9,0.975} \cdot \sqrt{\mathrm{var}_{rg}\left(\tilde{\theta}_j\right)} \tag{3}$$

where:

$\hat{\theta}_j = \dfrac{\hat{\bar{Z}}_{R,j} - \hat{\bar{Z}}_{T,j}}{\hat{\bar{Z}}_{T,j}}$ is the full sample estimate of the relative nonresponse bias on expenditure category j;

$t_{9,0.975}$ is the 97.5$^{\text{th}}$ percentile of a t-distribution with 9 degrees of freedom; and

$\text{var}_{rg}(\tilde{\theta}_j)$ is the random groups variance estimator described above.

2.6. *Categorical Data*

We compared respondents and proxy nonrespondents on categorical socio-demographic characteristics. For these comparisons, the test statistic was the adjusted Rao-Scott chi-square, a modified version of the Pearson chi-square, which accounts for the complex sample design (SAS Institute Inc. 2004). The null hypothesis for the two-way comparisons in our analyses is that there is no association between response status and subgroup. The null hypothesis for the two-way comparisons in our analyses is that there is no association between response status and subgroup. The null hypothesis for the one-way comparisons is that the respondent distribution for a characteristic from the CE is statistically equivalent to the corresponding distribution for the population obtained from the ACS.

3. **Individual Studies**

3.1. *Comparison of Response Rates Across Subgroups*

This study examined the response rates among socio-demographic subgroups that could be identified for both respondents and nonrespondents. The goal was to determine whether the survey's respondents and nonrespondents had the same socio-demographic characteristics. The subgroups analyzed were: region of the country (Northeast, Midwest, South, West), 'urbanicity' (urban, rural), type of PSU, housing tenure (owner or renter), and housing values for owners and renters.

Base-weighted response rates were calculated for these subgroups separately for Waves 1, 2, and 5 of the survey. They answer the question "What percent of the survey's target population do the respondents represent?" Base-weighted response rates are defined as the sum of base-weighted interviewed units divided by the sum of base-weighted interviewed units plus the units with Type A noninterviews.

$$\text{Base-weighted response rate} = \frac{\sum_i w_i I_i}{\sum_i w_i \left(I_i + A_i \right)} \tag{4}$$

where:

w_i = base weight for the ith consumer unit (CU);

I_i = 1 if the ith CU is a completed interview, and 0 otherwise; and

A_i = 1 if the ith CU is a Type A noninterview, and 0 otherwise.

Ideally, for this type of analysis, there should be no missing values for the subgroups of interest. Missing values may distort the response rates within a subgroup, especially if the cases with missing values differ from the complete cases. Therefore, this analysis was restricted to subgroups with no missing values for all the eligible cases. The available variables were primarily geographic frame variables. Most likely, other variables also contribute to response propensity.

Table 2 shows the weighted response rates for the subgroups in Waves 1, 2, and 5. For each subgroup, the sample size and response rate are given by wave. The results show that response rates differ within all of the subgroups examined. In particular, statistically significant differences ($p<0.05$) were found in the following pairwise comparisons within the subgroups:

- across the regions, CUs in the Northeast and West have lower response rates than those in the Midwest and South;

- across the types of PSU, CUs in metropolitan Core Based Statistical Areas (CBSAs) with a population of more than 2 million people have lower response rates than those in other types of PSUs;

- renters in the third and fourth housing value quartiles have lower response rates than renters in the first and second quartiles in the unit and area frames, with a similar trend among homeowners; and

- CUs in urban areas have lower response rates than those in rural areas.

In general, response rate differences within the subgroups suggest that the data are not MCAR because the respondent and nonrespondent CUs are not simple cross sections of the original sample.

Although there is evidence of an association between housing tenure and survey participation, there were no statistically significant differences between the pair-wise comparisons of owners and renters in Waves 2 and 5. Respondents who do not own or rent their homes had significantly higher response rates, though the number of these 'other' respondents is very small. These findings generally held across Waves 1, 2, and 5.

Table 2. Subgroup response rates by wave, CE

Subgroup	Wave 1		Wave 2		Wave 5	
	n	Response Rate %	n	Response Rate %	n	Response Rate %
Overall	12,451	78.9	12,372	76.4	12,505	76.5
Region[1,2,5]						
Northeast	2,468	75.1	2,468	74.8	2,332	73.8
Midwest	2,871	81.2	2,841	78.6	2,875	78.0
South	4,116	80.3	4,086	77.1	4,038	79.4
West	2,996	77.4	2,977	74.5	3,260	72.8
Type of PSU[1,2,5]						
A Metropolitan CBSAs > 2 million people	6,197	75.7	6,127	73.5	6,462	73.4
X Metropolitan CBSAs < 2 million	4,326	81.2	4,355	78.9	4,135	79.6
Y Micropolitan CBSAs	1,308	83.1	1,286	79.6	1,304	79.6
Z Non-CBSA ('rural') areas	620	80.9	604	78.0	604	77.0
Unit and Area Frames						
Housing value - Renters[1,2,5]						
Quartile 1-2	1,947	80.4	1,901	79.2	1,241	78.4
Quartile 3-4	1,891	76.0	1,910	75.0	1,171	74.6
Housing value - Homeowners[2,5]						
Quartile 1-2	4,228	79.8	4,239	77.2	2,869	78.3
Quartile 3-4	3,399	78.7	3,406	75.2	2,191	75.1
Urbanicity[1,2,5]						
Urban	10,217	78.2	10,159	75.9	10,054	75.7
Rural	2,234	81.4	2,213	78.6	2,451	79.4
Housing tenure[*2,5]*						
Owner	6,446	--	8,395	75.8	8,515	76.2
Renter	3,071	--	3,872	77.3	3,908	77.0
Other	93	--	105	92.6	82	88.4

[1, 2, 5] Indicates a statistically significant difference ($p<0.05$) was found for the computed Rao-Scott adjusted chi-square statistic for the test of no association between survey participation and subgroup in Waves 1, 2, and 5, respectively.

* Wave 1 response rates for housing tenure are not displayed due to missing values.

3.2. Comparison of CE Respondents to External Data

Another common approach to analyzing nonresponse is to compare the distribution of socio-demographic characteristics of respondents to that of a recent census or other 'gold standard' survey (Groves 2006). A 'gold standard' survey is one whose estimates are considered very accurate. Any significant differences between CE and the census or 'gold standard' survey suggest that respondents in CE are not representative of the target population, and thus the missing data in the CE are probably not MCAR.

The 'gold standard' survey chosen for this study was the 2005 American Community Survey (ACS). The ACS satisfied three important criteria: its estimates are considered to be very accurate; it has key socio-demographic variables available; and it was conducted in a time period very close to that which was used to analyze the CE. The ACS is a mandatory survey with a response rate of 97% and a coverage rate of 95% (Census Bureau 2006). The ACS data used in this study were obtained from published tables on the Census website.

CE data were weighted three ways for comparisons with the ACS: with base weights, noninterview adjustment weights, and final calibration weights. Comparisons were made for all three stages of weighting. Since all of CE's weights are CU-level weights and all of the ACS's weights are person-level weights, we made the results of the two surveys comparable by deriving frequency distributions for person-level characteristics in the CE from CU-level weights.

Table 3 presents a comparison of the distribution of selected respondent socio-demographic characteristics between CE and the ACS. The variables compared were: gender, age, race, educational attainment, household size, tenure, the number of rooms in the dwelling unit, housing value, rent, and CU income. Statistically significant differences ($p<0.05$) were found between the two distributions for all comparisons and all types of weighting with only two exceptions: calibration-weighted age and housing.

In addition to the different respondent characteristics, several other factors may contribute to making the differences statistically significant. First, the extremely large sample sizes make statistical significance likely even if the differences are not substantively meaningful. Second, CE and the ACS collect data differently. The two surveys use different data collection modes, and they use different question wording. As a result, the validity of the comparison to the ACS is limited by the extent to which the survey designs are truly comparable. Third, each CU was treated independently even though a CU can appear multiple times in the data if the CU participates in more

than one wave of the survey. This also has the potential to artificially magnify the differences between the demographic distributions of CE and the ACS.

The majority of the percentages shown in Table 3 had differences smaller than six percentage points, meaning that CE and ACS distributions are very similar on these variables. However, larger differences were found for race and rent: there were higher percentages of whites and higher percentages of monthly rents under $500 among Interview Survey respondents than among ACS respondents, indicating that the Interview data are probably not MCAR.

In short, the first study found that the data are not MCAR, and this study provided further evidence to substantiate that conclusion.

Table 3. A comparison of selected characteristics of respondents in the CE and the American Community Survey

		CE Interview						CE Interview		
	ACS	Base-Weighted	Noninterview-Weighted	Calibration-Weighted			ACS	Base-Weighted	Noninterview-Weighted	Calibration-Weighted
Gender (%) [1,2,3]						*Number of rooms in dwelling unit (%)* [1,2,3]				
Male	49.0	48.2	48.2	48.2		1	1.0	1.0	1.0	1.1
Female	51.0	51.8	51.8	51.8		2	3.2	2.0	2.0	2.1
Age (%) [1,2]						3-4	26.9	22.8	23.0	23.6
Under age 25	34.5	35.4	34.4	34.8		5-6	40.8	39.5	39.4	39.5
25-34	13.5	12.4	12.2	13.4		7-8	20.2	24.0	23.9	23.5
35-44	15.0	14.7	14.6	14.7		9+	8.0	10.7	10.6	10.3
45-54	14.6	14.3	14.5	14.5		*Owner occupied housing value (%)* [1,2,3]				
55-64	10.5	10.7	11.1	10.5		<$50,000	9.1	7.5	7.4	7.6
65-74	6.4	6.7	7.1	6.3		$50,000 to $99,999	18.1	16.2	16.3	16.5
75 and over	5.7	5.8	6.1	5.7		$100,000 to $149,999	17.4	16.6	16.7	16.8
Race (%) [1,2,3]						$150,000 to $199,999	13.7	14.2	14.4	14.4
White	74.8	82.3	83.3	81.4		$200,000 to $299,999	15.1	17.2	17.0	17.0
Black	12.5	11.4	10.4	12.5		$300,000 to $499,999	15.4	16.6	16.6	16.5
Other	12.7	6.3	6.3	6.1		$500,000 to $999,999	9.3	9.6	9.5	9.3
Education attainment* (%) [1,2,3]						$1,000,000+	2.0	2.1	2.1	2.0
Less than high school	15.8	16.3	16.1	16.0		*Monthly rent (%)* [1,2,3]				
High school graduate	29.6	28.1	28.0	28.0		Less than $500	20.5	38.7	38.5	38.4
Some college/Assoc degree	27.5	28.1	28.1	28.3		500- <750	29.1	29.4	29.4	29.6
College graduate	27.2	27.4	27.7	27.7		750- <1,000	22.0	16.1	16.4	16.4
Household size (%) [1,2,3]						1,000- <1,500	16.1	9.8	9.9	9.8
1 person	27.1	28.2	29.3	29.4		1,500+	6.3	4.2	4.3	4.2
2 persons	33.3	31.9	32.6	31.9		*CU income (%)* [1,2,3]				
3 persons	16.0	15.9	15.2	15.6		<$10,000	8.7	10.3	10.2	10.4
4+ persons	23.6	24.1	22.9	23.1		$10,000 to $14,999	6.2	7.6	7.6	7.6
Housing tenure (%) [1,2]						$15,000 to $24,999	12.0	12.6	12.6	12.6
Owner	66.9	68.0	68.8	67.4		$25,000 to $34,999	11.5	14.3	14.3	14.3
Renter	33.1	32.0	31.2	32.6		$35,000 to $49,999	15.1	15.1	15.2	15.2
						$50,000 to $74,999	18.9	16.4	16.4	16.5
						$75,000+	27.7	23.7	23.7	23.4

[1,2,3] Indicates a statistically significant difference (p<0.05) was found for the computed Rao-Scott adjusted chi-square statistic for the test of no difference in distributions between the ACS and the CE base-weighted, noninterview-weighted, and calibration-weighted, respectively

* respondent age ≥ 25

3.3 'Harder-to-Contact' Respondents as Proxies for Nonrespondents

The third study uses 'harder-to-contact' respondents as proxy nonrespondents. It draws on a theory known as the 'continuum of resistance' to identify appropriate respondents to serve as proxy nonrespondents. This theory suggests that sample units can be ordered across a continuum by the amount of interviewer effort exerted in order to obtain a completed interview (Groves 2006). This approach is limited by the extent that the proxy nonrespondents are like actual nonrespondents. Other studies (Lin and Schaeffer 1995) have suggested that there are significant differences between 'harder-to-contact' respondents and nonrespondents.

In past research, 'harder-to-contact' respondents were classified along the 'continuum of resistance' using the total number of contact attempts, the length of time the sample unit was in the field, and the last five percent of completed interviews (Bates and Creighton 2000; Traugott 1987; Lin and Schaeffer 1995; Ellis, Endo and Armer 1970). The weakness with these indicators is in attributing the contact difficulty to the correct party, the interviewer or the respondent. For instance, the interviewer's schedule or personal preferences for when to work on a case may affect the length of time the sample unit is in the field. Also, the total number of attempts may include scheduling appointments and other preliminary activities prior to actually attempting an interview.

Using data collected in the CHI, we defined respondents to be 'harder-to-contact' when over 45 percent of the contact attempts resulted in noncontacts. This cut-off for 'hard to contact' respondents was selected to yield a response rate slightly under 80 percent, which is similar to CE's actual response rate during the time period covered by the data. Also, this measure controls for some of the weaknesses cited above because it standardizes the amount of effort exerted by an interviewer to make contact across all sample units.

As an example, consider the contact history of a CU that had 6 contact attempts:

Contact attempt	Classification
1. No one home	Noncontact
2. No one home	Noncontact
3. Got answering machine/service	Noncontact
4. No one home	Noncontact
5. Respondent too busy, appointment set	Contact
6. Complete case – ready to transmit	Contact

In this example, 6 contact attempts were made, resulting in 2 contacts and 4 noncontacts, for a 67% noncontact rate. Since the noncontact rate is greater than 45%, this particular CU was classified as 'harder-to-contact.'

In this study respondents and proxy nonrespondents were compared at each wave of the survey on the following socio-demographic characteristics: gender, marital status, race/ethnicity, age, educational attainment, household tenure, Census Region (Northeast, Midwest, South and West), urbanicity (urban or rural) and CU size. Table 4 shows these comparisons. For example, in Wave 1 people who were 25-34 years old made up 15.1% of the respondents and 19.9% of the proxy nonrespondents. Differences such as this suggest that the data may not be MCAR with respect to some socio-demographic characteristics (respondent age, race, educational attainment, marital status, CU size, household tenure, and census region).

Table 4. Demographic characteristics of respondents and proxy nonrespondents ('harder-to-contact') in CE by wave

Demographic Characteristic	Wave 1		Wave 2		Wave 5	
	Respondents	Proxies	Respondents	Proxies	Respondents	Proxies
Gender (%)						
Male	41.7	41.5	40.0	42.0	38.8	38.8
Female	58.3	58.5	60.0	58.0	61.2	61.2
Age (%)[1, 2, 5]						
Under age 25	9.8	8.6	7.1	7.6	7.3	7.2
25-34	15.1	19.9	15.4	19.8	14.2	18.3
35-44	19.3	21.7	20.2	23.1	20.0	22.4
45-54	19.4	21.1	19.7	20.5	20.3	21.6
55-64	15.9	15.1	16.0	15.8	16.4	14.8
65-74	10.5	7.2	11.1	7.5	11.3	8.4
75 and over	10.0	6.3	10.5	5.7	10.4	7.3
Race (%)[1]						
White	83.4	82.6	83.4	83.3	82.6	82.7
Black	11.1	12.6	11.2	11.5	11.4	12.3
Other	5.5	4.8	5.4	5.2	6.0	5.0
Educational Attainment (%)[1, 2]						
Less than high school	15.1	14.1	15.5	13.1	15.3	13.9
High school graduate	25.7	23.9	26.1	25.1	26.1	27.2
Some college or Associate's degree	32.1	32.0	31.2	33.4	32.0	30.2
College graduate	27.2	30.1	27.2	28.3	26.6	28.7
Marital Status (%)[1, 2, 5]						
Not Married	45.1	52.0	45.6	50.5	45.4	49.4
Married	54.9	48.0	54.4	49.5	54.6	50.6
CU Size (%)[1, 2, 5]						
1	27.5	31.9	27.8	29.9	27.7	29.8
2	32.3	30.6	31.9	31.3	32.2	31.2
3	15.7	16.4	15.1	17.0	15.3	18.3
4+	24.5	21.1	25.1	21.8	24.8	20.6
Household Tenure (%)[2]						
Owner	68.4	66.3	68.8	64.9	68.4	67.9
Renter	31.6	33.7	31.2	35.1	31.6	32.1
CU Income (%) (imputed)						
<$10,000	-	-	9.7	8.6	8.5	8.2
$10,000 to $14,999	-	-	6.8	5.8	7.8	6.1
$15,000 to $24,999	-	-	13.3	11.3	12.4	12.0
$25,000 to $34,999	-	-	12.5	13.1	12.0	14.6
$35,000 to $49,999	-	-	13.7	14.1	15.5	15.9
$50,000 to $74,999	-	-	16.9	19.2	17.4	17.6
$75,000+	-	-	27.1	27.9	26.3	25.6
Census Region (%)[1, 2, 5]						
Northeast	18.6	17.1	18.4	20.3	18.5	17.9
Midwest	24.6	24.9	24.5	24.9	23.7	25.2
South	34.1	39.4	34.9	36.4	36.0	39.5
West	22.7	18.6	22.2	18.4	21.8	17.4
Urbanicity (%)						
Rural	20.9	21.2	20.9	20.8	23.2	24.3
Urban	79.1	78.8	79.1	79.2	76.8	75.7

[1, 2, 5] Indicates a statistically significant difference (p<0.05) between the respondents and nonrespondents for the demographic characteristic at the particular interview (1, 2, or 5)

In addition to answering the MCAR question, this study also estimated the relative nonresponse bias for total expenditures and for 13 expenditure subcategories: alcoholic beverages, apparel and services, cash contributions, education, entertainment, food, health care, housing, personal care, personal insurance and pensions, reading, tobacco and smoking supplies, and transportation. See Appendix A for a description of the expenditure categories. Table 5 shows the relative nonresponse bias estimates with their 95% confidence intervals for Waves 2 and 5.

When 'harder-to-contact' respondents were used as proxy nonrespondents, no evidence of nonresponse bias was found for total expenditures. The relative nonresponse bias for total expenditures was similar for both waves:

 -0.14% with a 95% confidence interval of (-1.40%, 1.12%) for Wave 2 and -0.10% with a 95% confidence interval of (-1.18%, 0.98%) for Wave 5. The nonresponse bias was not significant for most of the subcategories as indicated by the inclusion of zero in the confidence interval. The exceptions were reading materials and health expenditures in both Waves 2 and 5. Reading materials had a nonresponse bias of 3.82% (0.51%, 7.13%) for Wave 2 and 3.41% (0.28%, 6.54%) for Wave 5. The magnitude of nonresponse bias is similar for both waves. The magnitude of the bias was similar for both waves for health expenditures: 3.68% (1.80%, 5.56%) for Wave 2 and 3.21% (0.88%, 5.54%) for Wave 5. It is worth noting that these two categories represent only 0.29% and 6% of total spending, so their impact on total nonresponse bias is very small.

Table 5: Estimates of relative nonresponse bias in CE expenditure categories using harder to contact respondents as proxy nonrespondents

Expenditure Category	Share of total expenditures (%)*	Harder-to-contact respondents as proxy nonrespondents					
		Wave 2			Wave 5		
		Relative Bias %	Lower 95% CI	Upper 95% CI	Relative Bias %	Lower 95% CI	Upper 95% CI
Total Expenditures		-0.14	-1.40	1.12	-0.10	-1.18	0.98
Alcoholic Beverages	0.81	-0.96	-3.74	1.82	-2.82	-6.18	0.54
Apparel and Services	2.88	0.96	-3.48	5.40	0.37	-1.85	2.59
Cash Contributions	3.90	2.09	-2.60	6.78	0.73	-2.02	3.48
Education	1.90	-3.74	-10.83	3.35	1.86	-6.08	9.80
Entertainment	5.10	0.49	-2.49	3.47	0.40	-2.43	3.23
Food	13.40	0.19	-0.83	1.21	0.76	-0.06	1.58
Health	6.00	**3.68**	**1.80**	**5.56**	**3.21**	**0.88**	**5.54**
Housing	33.13	-0.28	-1.40	0.84	-0.89	-1.94	0.16
Personal Care	0.63	0.32	-2.24	2.88	0.59	-0.91	2.09
Personal Insurance	10.32	-0.51	-2.72	1.70	-0.68	-2.36	1.00
Reading Materials	0.29	**3.82**	**0.51**	**7.13**	**3.41**	**0.28**	**6.54**
Tobacco	0.72	-0.27	-4.07	3.53	-1.51	-3.43	0.41
Transportation	19.12	-1.96	-5.04	1.12	-0.53	-3.59	2.53

* Note that this column does not sum to 100% since some expenditure items are not listed in one of these subcategories.

3.4. Pattern of Participation while in Sample

The fourth study is based on the premise that we can learn about nonrespondents by examining panel survey respondents who failed to complete the entire series of interviews (Reyes-Morales 2003 and 2007). In this study, the proxy nonrespondents are attritors (survey dropouts) and intermittent respondents. They were compared to complete respondents on socio-demographic variables and expenditures.

The study was based on a single cohort of data, CUs that had their first interview in April-June 2005, their second interview in July-September 2005, and their last (5th) interview in April-June 2006. The cohort had 3,071 unique

CUs out of which 2,468 were used in the study. [1] Using the common data file described earlier, demographic characteristics were imputed for the nonrespondents with a technique similar to the 'last observation carried forward' (Verbeke and Molenberghs 2000). In this technique, any missing values for a particular CU were imputed by copying the values that were recorded for that CU in a previous interview. The technique assumes that the demographic characteristics do not change from one wave to the next.

The set of 2,468 usable CUs was divided into three groups according to their pattern of participation in Waves 2 through 5 while they were in the sample:

1. Complete respondents: These are CUs that participated in the survey in all contiguous waves of the survey period.

2. Attritors: These are CUs that participated in the first wave for which they were eligible, and possibly completed the second and third waves (if eligible) but then refused to participate in all subsequent waves for which they were eligible. For example, if the pattern of participation was (1, 2, 2, 3) or (1, 1, 2, 2), then that CU would be classified as an attritor. [2]

3. Intermittent respondents: These are CUs that participated in at least one wave among all the waves for which they were eligible. For example, if the pattern of participation was (3, 2, 1, 2), then that CU would be classified as an intermittent respondent.

The set of complete respondents had 1,941 CUs (78.6%), the set of intermittent respondents had 347 CUs (14.1%), and the set of attritors had 180 CUs (7.3%).

After accounting for CE's complex sampling design, the three response groups were compared using the Rao-Scott chi-square and Wald statistics on various socio-demographic variables, including household tenure (owner or renter), marital status, gender, respondent age, race, Hispanic origin, CU size, educational attainment, region, and urbanicity. Statistically significant differences were found between intermittent respondents and complete respondents with respect to age and Hispanic origin, while statistically significant differences were found between

---

[1] The 603 CUs excluded from this study were excluded for the following reasons: 332 CUs did not participate in the survey or did not provide enough information to count as a completed interview; 231 CUs dropped out of the survey after the first interview; and 40 CUs were ineligible for an interview.
[2] Here '1' indicates a completed interview, '2' indicates a Type A noninterview, and '3' indicates ineligibility for the sample.

attritors and complete respondents only with respect to age (Table 6). The difference in sample size between attritors and intermittent respondents, may have resulted in a loss of statistical power to detect demographic differences. This may be the cause of the lack of statistical difference between attritiors and respondents , rather than substantive differences between the groups.

For relative nonresponse bias computations, we combined attritors and intermittent respondents to form one group of proxy nonrespondents and averaged expenditures across Waves 2 through 5 for each expenditure category (Table 7).

When attritors and intermittent respondents were used as proxy nonrespondents, the relative nonresponse bias for total expenditures was not statistically different from zero, -0.54% (-2.31%, 1.24%). However, three expenditure subcategories showed some evidence of nonresponse bias: entertainment, 3.46% (0.35%, 6.57%), personal insurance, 3.82% (1.72%, 5.93%) and transportation, -5.65% (-9.11%, --2.19%). These subcategories represent 5%, 10%, and 19%, respectively, of total expenditures and are different from the expenditure subcategories identified in the 'harder-to-contact' study.

Table 6. Demographic characteristics of respondents and proxy nonrespondents (attritors and intermittent respondents) in CE

Demographic Characteristic	Respondents (n=1,941)	Proxy nonrespondents Intermittent (n=347)	Attritors (n=180)		Demographic Characteristic	Respondents (n=1,941)	Proxy nonrespondents Intermittent (n=347)	Attritors (n=180)
Gender (%)					*CU Size (%)*			
Male	41.1	46.1	38.9		1 person	30.6	25.6	35.0
Female	58.9	53.9	61.1		2 persons	31.1	30.3	29.4
Age (%)[I,A]					3 persons	15.1	15.0	16.1
Under age 25	7.8	6.2	9.5		4+ persons	23.2	29.1	19.4
25-34	16.8	17.2	22.9		*Household Tenure (%)*			
35-44	20.8	25.1	24.0		Owned	62.7	64.3	56.1
45-54	19.9	21.9	21.8		Rented	35.4	34.9	41.7
55-64	15.4	17.5	9.5		Other	1.9	0.9	2.2
65-74	10.6	6.5	7.3		*CU Income(%) (imputed)*			
75 and over	8.6	5.6	5.0		<$10,000	9.6	8.4	11.7
Race (%)					$10,000 to $14,999	7.4	7.2	4.4
White	81.6	78.7	80.6		$15,000 to $24,999	12.9	10.4	16.1
Black	11.8	14.7	13.3		$25,000 to $34,999	12.3	13.6	10.6
Other	6.6	6.6	6.1		$35,000 to $49,999	15.4	14.2	22.2
Hispanic Origin (%)[I]					$50,000 to $74,999	17.6	18.2	15.0
Hispanic	11.3	17.3	10.6		$75,000+	24.8	28.0	20.0
Non-Hispanic	88.7	82.7	89.4		*Census region*			
Educational Attainment (%)					Northeast	18.6	16.4	22.8
Less than high school	15.3	16.0	15.2		Midwest	23.7	21.3	21.7
High School graduate	26.2	25.6	24.7		South	34.0	38.9	29.4
Some college or Associate's degree	32.8	31.4	38.2		West	23.7	23.3	26.1
Bachelor's or higher	25.7	27.0	21.9		*Urbanicity (%)*			
Marital Status (%)					Urban	82.5	84.4	86.1
Married	51.1	54.5	45.0		Rural	17.5	15.6	13.9
Not married	48.9	45.5	55.0					

'I' indicates a statistically significant difference (p<0.05) between complete respondents and intermittent respondents, and 'A' between complete respondents and attritors.

Table 7. Estimates of relative nonresponse bias in CE expenditure categories using intermittent respondents and attritors as proxies for nonrespondents

Expenditure Category	Share of total expenditures (%)*	Intermittent respondents and attritors as proxy nonrespondents		
		Across waves		
		Relative Bias %	Lower 95% CI	Upper 95% CI
Total Expenditures		-0.54	-2.31	1.24
Alcoholic Beverages	0.81	-2.37	-9.09	4.36
Apparel and Services	2.88	-0.01	-5.24	5.21
Cash Contributions	3.90	4.24	-0.98	9.46
Education	1.90	-2.27	-11.81	7.28
Entertainment	5.10	**3.46**	**0.35**	**6.57**
Food	13.40	0.09	-1.35	1.53
Health	6.00	1.03	-2.76	4.82
Housing	33.13	-0.56	-2.26	1.14
Personal Care	0.63	1.95	-0.74	4.65
Personal Insurance	10.32	**3.82**	**1.72**	**5.93**
Reading Materials	0.29	4.22	-1.16	9.60
Tobacco	0.72	2.51	-1.79	6.81
Transportation	19.12	**-5.65**	**-9.11**	**-2.19**

* Note that this column does not sum to 100% since some expenditure items are not listed in one of these subcategories.

4. Findings Across Studies

4.1. MCAR

Using socio-demographic data from respondents and nonrespondents, all four studies conclude that either the respondents are significantly different from the complete sample or that they are significantly different from the nonrespondents. In either case, these findings suggest that the data are not MCAR. (See Tables 2, 3, 4, and 6.)

4.2 Characteristics of Nonrespondents

All four studies provide insights into the characteristics of CE nonrespondents. Table 8 synthesizes the results from the four studies. In Table 8, a characteristic is shown as being under-represented (or over-represented) if that characteristic was under-represented (or over-represented) in at least three of the four studies. For example, race was analyzed in three studies (the comparison to external data, and the two proxy nonrespondent studies), and all three of them found blacks to be under-represented (see Tables 3, 4, and 6). Therefore, we conclude that blacks are more likely than other races to be CE nonrespondents. Likewise, age was analyzed in three studies, and all three of them found people in the oldest age group to be over-represented, so we conclude that people over 55 years old are more likely than people in younger age groups to be CE respondents.

4.3. Estimates of Relative Nonresponse Bias for Expenditure Categories

Of the four studies, only the two proxy nonrespondent studies estimate nonresponse bias. Both proxy nonrespondent studies found no evidence of nonresponse bias in the total expenditures category, and no evidence of nonresponse bias in most of the thirteen expenditure subcategories examined. Although both studies found some evidence of nonresponse bias in a few of the expenditure subcategories, they did not find it in the same subcategories. Since the two studies defined 'proxy' nonrespondents differently, some variation in the results was expected, but taken together the results suggest that nonresponse bias is probably not a major issue in the CE (see Tables 5 and 7).

Table 8. Summary of characteristics of respondents across studies

Characteristic	Response Rate across subgroups	Comparison to External Data	'Harder-to-contact' as nonrespondents	Attritors & intermittent respondents as nonrespondents	Conclusion
Gender	—	males under-represented	inconclusive	inconclusive	inconclusive
Age	—	55+ over-represented	55+ over-represented	65+ over-represented	55+ over-represented
Race	—	blacks under-represented	blacks under-represented	blacks under-represented	blacks under-represented
Educational attainment	—	HS graduate, some college & higher over-represented	HS graduate or less over-represented	HS graduate over-represented	inconclusive
Housing tenure	renters over-represented	renters under-represented	renters under-represented	inconclusive	inconclusive
Marital Status	—	—	not married under-represented	inconclusive	inconclusive
Census region	Midwest over-represented	—	Midwest & South under-represented; West over-represented	Midwest over-represented	inconclusive
Urbanicity	urban under-represented	—	inconclusive	urban under-represented	inconclusive
CU size	—	1 person CU over-represented	2-person & 4+ person CUs over-represented	2-person CUs over-represented	inconclusive

— indicates characteristic was not studied

5. Conclusion

The CE is a federal statistical survey that has a response rate below 80 percent. Each of the four studies was designed to analyze nonresponse in CE by answering one or more of the following questions: (1) Are the data in the CE MCAR? (2) What are the demographic characteristics of the nonrespondents? and (3) What is the level of nonresponse bias in the CE? All four studies shed light on the first two questions, while the two proxy nonrespondent studies addressed the third question.

All of the studies found that the data are not MCAR. The study comparing response rates by subgroup addressed the MCAR question by looking for differences in response propensities on variables that were available for both respondents and nonrespondents. Statistically significant differences were found by region of the country, urbanicity, and housing tenure. Likewise, the study comparing respondent demographic characteristics to the American Community Survey's population found statistically significant differences for most of the variables examined. The study using 'harder-to-contact' respondents as proxy nonrespondents found statistically significant differences for age, marital status, CU size, and region of the country for all waves, and statistically significant differences for race, educational attainment, and household tenure for some waves. The study using intermittent respondents and attritors as proxy nonrespondents found statistically significant differences by age in both groups of proxy nonrespondents, and for Hispanic origin in the set of 'intermittent' proxy nonrespondents. Because statistically significant differences were found in each of these studies, we conclude that the data are not MCAR.

All four of these studies addressed the second question by providing insight into the characteristics of CE nonrespondents. Any characteristic for which a statistically significant difference was observed suggests that the respondent sample over-represents particular subgroups of the survey's target population while under-representing other subgroups. Again, the individual study conclusions vary, but it is evident that blacks are under-represented among the respondents (nonrespondents tend to be disproportionately black) while those age 55 and over tend to be over-represented (respondents tend to be disproportionately older).

Two of the studies addressed the third question about the level of nonresponse bias in the survey's estimates. The study that used 'harder-to-contact' respondents as proxy nonrespondents did not find any nonresponse bias in the total expenditures category, since its 95% confidence intervals included the number zero. In addition, the study did not find any nonresponse bias in eleven of the thirteen expenditure subcategories examined. The results from the

'patterns of participation' study were similar. That study did not find any nonresponse bias in the total expenditures category, and it did not find any nonresponse bias in ten of the thirteen expenditure subcategories examined.

Although both proxy nonrespondent studies identified a few expenditure categories with non-zero levels of nonresponse bias, they identified different categories. Since some bias could be expected to occur at random, and since the findings across the two proxy studies were not consistent, we conclude that CE expenditure estimates are not subject to high levels of nonresponse bias.

No one study, taken alone, is meant to provide a definitive answer to the questions raised in this research. Each approach has its own strengths and weaknesses. But taken together, the four studies show that CE estimates do not have a significant amount of nonresponse bias, even though the respondents and nonrespondents have different characteristics and the data are not MCAR.

The results from these four studies support recent research on nonresponse bias that indicates nonresponse rates are a weak predictor of nonresponse bias and only indicate a risk for nonresponse bias. Research indicates that everyone has an unobservable propensity to be either a respondent or nonrespondent in a survey (Groves et al. 2006; Groves and Peytcheva, 2008). Other sources of bias and measurement error exist in surveys and high response rates may not be adequate to prevent them.

These studies provide a counterexample to the commonly held belief that if a survey's data are not MCAR then its estimates are subject to nonresponse bias. From the nonresponse bias equation, the nonresponse bias in a survey estimate will be zero if there is complete response, or if the mean expenditure for respondents is equal to the mean expenditure for nonrespondents. The key for CE is the latter situation. When the data are not MCAR, as in this case, similar expenditure patterns for the respondents and nonrespondents could still be observed if the bias in an under-represented group (e.g., blacks) is counterbalanced by the same bias in an over-represented group (e.g., the over-55 age group). The absence of meaningful bias in the total expenditures category seems to support this argument. In conclusion, nonresponse bias does not appear to be a significant issue for the Consumer Expenditure Interview Survey.

6. References

Allison, P. (2002)**.** Missing Data. Sage University Series on Quantitative Applications in the Social Sciences**.** Thousand Oaks, CA: Sage.

Bates, N. and Creighton, K. (2000). The Last Five Percent: What Can We Learn From Difficult/Late Interviews? Proceedings of the American Statistical Association, Section on Government Statistics and Section on Social Statistics, 120-125.

Brick, J. M., and Bose, *J*. (2001). Analysis of Potential Nonresponse Bias. Proceedings of the Annual Meetings of the American Statistical Association, [CD-ROM].

Census Bureau (2006). Using the Data: Quality Measures. Retrieved April 18, 2007 from Census Bureau Web Site www.census.gov/acs/www/acs-php/quality_measures_coverage_2005.php.

Dixon, J. (2001). Relationship Between Household Nonresponse, Demographics and the Unemployment Rate in the Current Population Survey. Proceedings of the American Statistical Association, Section on Survey Methods Research, 803-806.

Dixon, J. (2004). Using Census Match Data to Evaluate Models of Survey Nonresponse. Proceedings of the American Statistical Association, Section on Survey Methods Research, 3421 -3425.

Duncan, G.J., and Hill, D.H. (1989). Assessing the Quality of Household Panel Data: The Case of the Panel Study of Income Dynamics. Journal of Business & Economic Statistics, *7*(4), 441-452.

Ellis, R.A., Endo, C.M. and Armer, J.M. (1970). The Use of Potential Nonrespondents for Studying Nonresponse Bias. The Pacific Sociological Review*, 13*(2), 103-109.

Groves, R.M. (2006). Nonresponse Rates and Nonresponse Bias in Household Surveys. Public Opinion Quarterly*, 70*(5), 646-675.

Groves, R.M., Couper, M.P., Presser, S., Singer, E., Tourangeau, R., Acosta, G., and Nelson, L. (2006). Experiments in Producing Nonresponse Bias. Public Opinion Quarterly*, 70*(5), 720-736.

Groves, R.M. and Peytcheva, E. (2008). The Impact of Nonresponse Rates on Nonresponse Bias. Public Opinion Quarterly, *72*(2), 167-189.

Lin, I.F, and Schaeffer, N. (1995). Using Survey Participants to Estimate the Impact of Nonparticipation. Public Opinion Quarterly, *59*(2), 236-258.

Little, R. J. A., and Rubin, D. B. (2002), Statistical Analysis with Missing Data, 2nd Edition, New York: John Wiley.

Pearl, D.K, and Fairley, D. (1985). Testing for the Potential for Nonresponse Bias in Sample Surveys. Public Opinion Quarterly, *49*(4), 553-560.

Purdie, D.M., Dunne, M.P., Boyle, F.M., Cook, M.D., and Najman, J.M. (2002). Health and Demographic Characteristics of Respondents in Australian National Sexuality Survey: Comparison with Population Norms. Journal of Epidemiology and Community Health*, 56,* 748-753.

Reyes-Morales, S.E. (2003). Characteristics of Complete and Intermittent Responders in the Consumer Expenditure Quarterly Interview Survey. Final Report, Bureau of Labor Statistics, Washington, D.C. .

Reyes-Morales, S.E. (2007). Characteristics of Complete and Intermittent Responders in the Consumer Expenditure Quarterly Interview Survey: 2003-2006. Final Report, Bureau of Labor Statistics, Washington, D.C. .

SAS Institute Inc. (2004). SAS/STAT® User's Guide, Version 9.1. Cary, NC: SAS Institute Inc.

Traugott, M.W. (1987). The Importance of Persistence in Respondent Selection for Preelection Surveys. Public Opinion Quarterly, *51*(1), 48-57.

Verbeke, G. and Molenberghs, G. (2000). Linear Mixed Models for Longitudinal Data. New York: Springer-Verlag.

Wolter, K.M. (1985). Introduction to Variance Estimation. New York: Springer-Verlag.

Appendix A. Expenditure Examples for the Aggregated Expenditure Categories

Aggregated Expenditure Group	Types of Expenditures
Alcoholic Beverages	Alcohol for home consumption plus alcohol at restaurants and bars.
Apparel and Services	Clothing, other apparel products, and footwear. Services including repair of shoes, watches, and jewelry, alterations, clothing rental, storage, and sewing materials.
Cash Contributions	Cash contributions to religious organizations, educational institutions, political organizations
Education	College, elementary, and high school tuition, books and supplies; recreational classes
Entertainment	Toys, games, arts, crafts, and other entertainment
Food	Food consumed at home and food consumed away from home (e.g., restaurants, take out and delivery, vending machines)
Health	Health insurance, physician, dental, and eye care services, hospital costs, prescription drugs, medical equipment
Housing	Owned and rented dwellings, mortgage interest, property taxes, maintenance, repairs, insurance, landscaping, vacation homes, lodging on out-of-town trips, utilities
Personal Care	Personal care products and services, electric personal care appliances, personal care services
Personal Insurance	Personal insurance, including life insurance and pensions, social security
Reading Materials	Newspaper, magazine, and books
Tobacco	Tobacco products and smoking supplies
Transportation	Vehicle purchases, gasoline and motor oil, vehicle finance charges, maintenance and repairs, vehicle audio equipment, vehicle insurance, rented vehicles, public transportation