Nonresponse in the American Time Use Survey[*]

Phawn M. Letourneau[1], Andrew A. Zbikowski[1]
[1]U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233

Abstract
The American Time Use Survey (ATUS) provides estimates of how people divide their time each day in the United States. Households that have completed their eighth and final monthly interview in the Current Population Survey (CPS) are eligible for selection in the ATUS. Since the initial data collection in 2003, the ATUS response rate has been below 60%. The low response rate has caused concern over the potential for bias in the ATUS estimates. In 2005, the Bureau of Labor Statistics conducted research on nonresponse in the 2004 ATUS. We follow with an analysis of nonresponse on the 2006 ATUS. We link the ATUS weighted output files to the call history files for nonresponse unit information, and link to the CPS files for available frame variables. We use SAS PROC SURVEYLOGISTIC to model response propensities in the ATUS for individuals or households with different characteristics.

Key Words: Survey; Nonresponse; Logistic Regression

1. Introduction

The American Time Use Survey (ATUS) is designed to provide estimates of how people spend their time each day in the United States. It examines how many hours Americans spend participating in activities such as work, childcare, religious services, and exercise. One use of the data is to compare time-use patterns among demographic subgroups. Since the initial data collection in 2003, the ATUS response rate has consistently been below 60%. The low response rate has caused concern over the potential for bias in the ATUS estimates. The Bureau of Labor Statistics (BLS) (Abraham et al., 2005) conducted research on nonresponse in the 2004 ATUS, examining which household and individual characteristics are related to survey nonresponse. This paper presents an analysis of nonresponse in the 2006 ATUS, with comparisons to earlier nonresponse data.

2. An Overview of the ATUS

The ATUS sample is drawn from the Current Population Survey (CPS). Households that have completed their eighth and final monthly interview in the CPS are eligible for selection in the ATUS. A sample size of approximately 2,000 households is randomly selected each month for interviewing. Households with minorities and households with children are oversampled to improve the reliability of the estimates for those subgroups. An eligible person age 15 years or over is randomly selected from each household to be the designated person who reports on activities for a designated day of the week. The sample is split evenly between weekdays and weekends, and across the weeks of the year. Thus, 10 percent of the sample is allocated to each weekday and 25 percent of the sample is allocated to each weekend day, due to interest in weekday/weekend comparisons.

Respondents are included in the ATUS two months after their household is retired from the CPS. Respondents are notified that they have been selected for the ATUS with an advance letter. The interview is conducted the day after the designated reporting day by the U.S. Census Bureau's computer-assisted telephone interviewing (CATI) center. A $40 incentive is provided to sample cases for which a phone number is not available on the CPS file. Efforts to contact an eligible sample person on the interview day may continue for up to 8 weeks.

3. Survey Nonresponse

Survey nonresponse refers to the failure to get a useable response from a potential respondent selected for the sample. Nonresponse may occur for several reasons, including noncontacts, refusals, or language barriers. Nonresponse is a source of bias in the survey estimates to the extent that: (1) those who respond are systematically different from those who do not respond with respect to the characteristic of interest; or (2) response propensities are

---

correlated with the estimates of interest. Thus, the degree of nonresponse bias is a function of how much the respondents and nonrespondents differ on the survey variables of interest, as well as the nonresponse rate.

We look at three different dimensions of response: (1) contact, (2) cooperation, and (3) overall response. We define these later in this paper.

4. Background for Research

Flanagan and Dexter (2007) conducted an analysis of nonresponse in the 2003 National Survey of College Graduates (NSCG). They used weighted logistic regression to examine the effect that various demographic characteristics had on nonresponse. They compared results to a previous study conducted on the 1993 NSCG and to other studies of the population in general. A number of variables were significantly related to nonresponse in the NSCG. These variables included age, marital status, race, citizenship, region, gender, occupation, employment, and disability. Higher nonresponse rates occurred for younger people, those who were separated, blacks, non-U.S. citizens, those from the West and South, males, Social Scientists and Non-Science and Engineering Occupations, those who were unemployed, and those who were disabled. Note that the NSCG population and the ATUS population are not the same.

Abraham, Maitland, and Bianchi (2005) conducted research on nonresponse in the 2004 ATUS, using weighted logistic regression to examine household and individual characteristics associated with survey nonresponse. They speculated that those who were busier or less socially integrated (e.g., more transient and less engaged in their communities) would be less likely to respond. Characteristics related to "busyness" and/or "social integration" included number of hours worked, presence of children in the home, marital status, employment status, housing tenure, and urbanicity. Gender, age, race-ethnicity, household income, education, and region were also considered. They found that people who were not in the labor force, separated or never married, living in a central city, or renting had lower response rates, supporting the "social integration" theory. People who worked longer hours had lower response rates due to noncontact. Those who were young, Hispanic or black, less educated, or had lower incomes also had lower response rates due to noncontact. Nonresponse was also high for people who refused to report their household income in the CPS.

Adams and Shoemaker (2004) performed logistic regression analysis, examining demographic and operational variables in the ATUS survey, to determine why response rates dropped between the pre-fielding period (November-December 2002) and the production phase (which began in January 2003). They found that cases missing the CPS family income variable were more likely to be ATUS refusals than cases where the CPS family income variable was not missing. They also found that younger people and black people were less likely to be contacted. Call patterns were also examined, and it was recommended that calls be spread more evenly throughout the day to increase contact rates.

5. Analysis Methodology

5.1 Distribution of Survey Outcome Codes
The ATUS response rates are calculated using the American Association for Public Opinion Research (AAPOR) response rate definition, RR2:

$$RR2 = \frac{C}{C + R + NC + O + UE} \qquad (1)$$

where:

C = completed interviews (including sufficient partial interviews)
R = refusals
NC = noncontacts
O = other noninterviews
UE = cases of unknown eligibility

These terms are defined from the final call outcome codes on the call history files (see ATUS data dictionary, 2007).

Contact rates are calculated using the AAPOR contact rate definition, CON1:

$$CON1 = \frac{C + R + O}{C + R + NC + O + UE} \qquad (2)$$

The noncontact rate is 1-*CON1*.

Cooperation rates are calculated using the AAPOR cooperation rate definition, COOP2:

$$COOP2 = \frac{C}{C + R + O} \qquad (3)$$

The response rate defined in (1) is the product of the contact rate in (2) and cooperation rate in (3). For example, for the total sample, the contact rate (80.25%) times the cooperation rate (71.76%) yields the response rate (57.58%).

5.2 Logistic Regression Analysis
We conduct an analysis of nonresponse in the 2006 ATUS to examine the effect that various characteristics have on nonresponse, and to compare our findings to the results described in section 4. Data are obtained by linking the ATUS weighted output files to the call history files for nonresponse unit information, and to the CPS microdata files for available frame variables. SAS PROC SURVEYLOGISTIC is used to model response propensities. Two advantages of analysis using weighted logistic regression are to determine the effect of a characteristic in the presence of other characteristics rather than in isolation, and to examine the effect of interactions. However, weighted logistic regression only partially adjusts for aspects of complex sample design. Using PROC SURVEYLOGISTIC purports to adjust for the stratified sample design in the ATUS, in which households are stratified based on the race-ethnicity of the householder, the presence and age of children in the household, and the number of adults in adult-only households.

The variables chosen for the primary analysis are shown in Tables 1 through 3. These are independent variables that may help explain the survey nonresponse. Logistic regression models are created for each of these groups. We analyze these for the contact model, the cooperation model, and the overall response model. Thus, the dependent variable is a binary variable indicating contact/noncontact, cooperation/ refusal, or response/nonresponse. The model formulation is:

$$P(\tilde{x}) = \frac{1}{1 + \exp(-\tilde{x}'\tilde{\beta})} \qquad (4)$$

where:

$P(\tilde{x}) =$ propensity to respond

$\tilde{x}'\tilde{\beta} = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k$, for k independent variables

The ATUS baseweights are used for each observation. They are adjusted to the entire year, since the ATUS weighting is processed quarterly. The program used for the analysis is the SURVEYLOGISTIC procedure in SAS, using the last level of each categorical variable as the reference level (see SAS OnlineDoc, 2003). Forward selection stepwise regression is used with a 0.01 level of significance. Available frame variables that show no correlation with nonresponse are removed from the analysis. Interactions among variables are checked, and no interactions remain in the final models.

Coefficient estimates, standard errors, and odds ratios are produced from the logistic regression runs. A coefficient that is significant and positive implies that having the characteristic in question raises the probability of being contacted by, cooperating with, or responding to the survey. A coefficient that is significant and negative implies that having the characteristic in question lowers the probability of being contacted by, cooperating with, or responding to the survey. The coefficients from the logistic regression are in terms of the log odds,

$$\log(odds) = \log[p/(1-p)] = bX \qquad (5)$$

where *p* is the probability of responding to the survey.   Therefore, the ratio of the odds can be computed by taking the exponential of the slope coefficient.

Comparisons from odds ratios are discussed in terms of relative effect. Statements made in the form of "category one is 20% more likely to respond than category two" means that if category two had a 0.05 probability of refusing, then category one would have a refusal probability of 1.20*0.05=0.06. Results of the logistic regression runs are shown in Tables 1 through 3.

6. Findings

6.1 Summary
Age, education, urbanicity, and income are significant across all three dimensions of contact, cooperation, and overall response. Other variables had some significance on one or two of the dimensions, such as housing tenure, gender, and region. Race-ethnicity and marital status did not show significant effects, contrary to earlier results in section 4. The findings for hours worked and presence of children were inconclusive as these variables showed significance for overall response, but not the dimensions of contact and cooperation.

6.2 Response
6.2.1 Age
The 66 and over age group has a higher likelihood of responding (65.82%) than age groups 30 and under (50.23%), 31-45 (55.41%), and 46-55 (59.55%). This is mostly driven by the high contact rate (92.95%) for the over 66 group, since they are more likely to be at home. It appears that persons under 30, or between the ages of 31 and 45, are over 4 times more likely to be a noncontact than a person over 65. Older people also have lower tendencies to refuse (16.73%), with the highest tendency to refuse in the 31 to 45 age group (22.83%). This group is 1.5 times more likely to refuse than the over 66 group.

6.2.2 Education
Those with a graduate degree have a higher likelihood of responding (70.26%) than those with less than a high school education (54.14%), a high school diploma (53.27%), some college (57.29%), and a bachelor's degree (63.37%). The differences are mostly due to the lower contact rates for those with less than a bachelor's degree, which fall around 78.00%. Those with less than a bachelor's degree are about 2 times more likely to be a noncontact than those with a graduate degree. Those with a graduate degree also have a higher likelihood of cooperating (78.87%), with the lowest tendency to cooperate in the high school education group (68.56%). They are 1.5 times more likely to be a refusal than those with a graduate degree.

6.2.3 Urbanicity
Respondents living in non-core based statistical rural areas have a higher likelihood of responding (63.36%) than those in core based statistical areas, central city (53.65%) and not in central city urban areas (57.96%). This is due to both a higher contact rate and a lower refusal rate for respondents living in non-core based statistical rural areas, 85.40% and 18.70%, respectively. Those living in core based statistical areas, central city and not in central city urban areas are about 1.5 times more likely to be a noncontact or a refusal. This supports the "social integration" theory discussed in section 4 (Abraham et al., 2005).

6.2.4 Income
Those with a missing value for income have a lower likelihood of responding (46.72%) relative to those with an income of $75,000 or more (65.34%). Much of this effect is due to the high refusal rate for those with a missing value for income (32.04%). Those with a missing value for income are 2 times more likely to be a refusal than those with an income greater than $75,000. In addition, the low-income group (less than $20,000) has a lower likelihood of being contacted (74.20%) than those with an income of $75,000 or more (86.48%). They are 1.5 times more likely to be a noncontact. These findings support earlier studies showing high nonresponse (refusals) for people who refused to report their income in the CPS (Adams and Shoemaker, 2004), and high nonresponse (noncontact) for those with lower incomes (Abraham et al., 2005).

6.2.5 Housing Tenure
Owners have a higher likelihood of responding (61.39%) than renters (48.42%). This is driven by contact rates (85.22% for owners versus 68.35% for renters). Owners are about 2 times more likely to be contacted than renters. This supports the "social integration" theory described in section 4.

6.2.6 Gender
Males have a slightly lower likelihood of responding (55.14%) than females (59.52%). This is due to cooperation rates (69.65% for males versus 73.40% for females). Males are 1.2 times more likely to be a refusal than females. This finding is consistent with the 2003 NSCG study (Flanagan and Dexter, 2007).

6.2.7 Region
Those living in the Midwest have a higher likelihood of responding (61.56%) than those living in the Northeast (55.42%), South (56.14%), and West (57.65%). This is because they are more likely to be contacted (83.04%). Those from the South are the least likely to be contacted (77.63%). They are 1.3 times more likely to be a noncontact than those living in the Midwest. This finding is consistent with the 2003 NSCG study.

6.2.8 Hours Worked and Presence of Children
People who work less than 35 hours a week (62.93%) have a slightly higher likelihood of responding, relative to those whose hours vary (54.03%). However, it is uncertain whether this is due to noncontact or refusal. Although this group has the highest cooperation rate (76.04%), and second to highest contact rate (82.76%), the effects are not significant for these models.

The results are also inconclusive for presence of children in the household. The overall model shows households with children are more likely to respond relative to households with no children, 2 or more adults. However, the response, contact, and refusal rates do not support this conclusion, nor are the effects significant for the contact and cooperation models. Note that this study uses different categories for children present in the home than the previous studies.

6.2.9 Race-ethnicity and Marital Status
Race-ethnicity and marital status did not show significant effects across the three dimensions of nonresponse, but a few estimates are mentioned here. White nonhispanics are the most likely to respond (61.61%), due to a high contact rate (84.33%) and cooperation rate (73.06%). Black nonhispanics are the least likely to respond (47.51%), due to a low contact rate (70.90%) and cooperation rate (67.01%). The widowed have the highest response rate (62.92%), due to a high contact rate (90.50%) and low refusal rate (18.09%); followed by married people. The never married group has the lowest response rate (51.26%), due to a low contact rate (72.55%) and high refusal rate (21.70%).

6.3 Results Similar to Studies in Section 4
Our findings are consistent with the following results taken from the earlier studies described in section 4:

- Lower response rates for people living in a central city and renters.
- Lower contact rates for people with less education, lower incomes, and in younger age groups.
- Higher refusal rates for people missing household income in the CPS.
- Higher response rates and contact rates for people living in the Midwest.
- Lower response rates and cooperation rates for males.

6.4 Results Differing from Studies in Section 4
We did not see significant effects in the 2006 ATUS for the following, as found in the previous studies described in section 4:

- Lower response rates for people who are unemployed or not in the labor force, and for people who are separated or never married.
- Lower contact rates for people who work longer hours, and for people who are hispanic or black.

Some of the differences in the results can be attributed to different methods of analysis, changes in populations, or different data collection methods. For example, BLS did not use logistic regression in the 2004 ATUS study; the NSCG study only targets college graduates; and data collection methods or call patterns for the ATUS may have changed from 2004 to 2006.

7. Discussion

7.1 Limitations

One limitation of the analysis is response bias in the frame. Since the ATUS sample is selected from the CPS, only CPS respondents make up the frame. This may bias the results if the CPS respondents are different than the CPS nonrespondents. There could also be item nonresponse in the frame, such as missing or imputed income for CPS. There is also the potential for month-in-sample (MIS) bias associated with the CPS MIS 8. The CPS has a high response rate, so the degree of bias may be limited.

A second limitation is we do not discuss the results of significance tests or implied statistical comparisons other than the coefficients for the logistic regression models. However, additional significance tests were performed to check the consistency of our logistic regression results. Ultimately, we were looking for patterns that may help explain the ATUS nonresponse rates.

A third limitation of the analysis is the logistic regression model assumes that the independent variables are linearly related to the log odds of the dependent "dichotomous" variable. If that linearity is not true, it could bias the results. Another limitation may come from interpreting the odds ratios. These are often used in epidemiology in statements like "twice as likely". However, they are only approximations to relative risk and may be overstated.

7.2 Recommendations

One result of the analysis may be a change to the unit nonresponse weighting adjustment procedure or item nonresponse imputation methods. The next step would be to compare estimates based on the current weighting procedure with those based on nonresponse weighting adjustments stemming from our response propensity models. Changes to operational procedures may also be considered, such as call patterns and refusal conversions. These procedures would apply to individuals who are harder to contact or more likely to refuse.

Another result of the analysis is the creation of procedures for doing response bias studies for surveys with frame variables using logistic regression via PROC SURVEYLOGISTIC. For example, this research is an effort for the Demographic Surveys Sample Redesign program, since we can use logistic regression for response bias studies for the 2010 redesign.

References

Abraham, Katherine G., Maitland, Aaron, and Bianchi, Suzanne (2005), "Nonresponse in the American Time Use Survey: Who is missing from the Data and How Much Does It Matter?," Presented at the American Time Use Survey Early Results Conference, Bethesda, Maryland, December 8-9, 2005.

Adams, Tamara and Shoemaker, Harland (2004), "ATUS Response Rates Analysis Results (ATUS-09)," Memorandum for Documentation, June 7, 2004.

Bureau of Labor Statistics and U.S. Census Bureau (2007), "American Time Use Survey User's Guide: Understanding ATUS 2003 to 2006," [http://www.bls.gov/tus/ atususersguide.pdf], December 2007.

Bureau of Labor Statistics and U.S. Census Bureau (2007), "American Time Use Survey Data Dictionary: 2006 Survey Methodology Data: Variables about ATUS,"   [http://www.bls.gov/tus/atussmcodebk06.pdf], June 2007.

Flanagan, Patrick E. and Dexter, Nancy (2007), "Nonresponse in Surveys of the College Educated, Revisited," Presented at the Joint Statistical Meeting, Salt Lake City, Utah, July 29-August 2, 2007.

SAS OnlineDoc, The SURVEYLOGISTIC Procedure, http://support.sas.com/onlinedoc/913/ docMainpage.jsp, 2003.

Table 1: Contact Model Results

Variable	Category	Parameter	Std. Error	Pr>ChiSq.	Odds Ratio
Gender	Male	Not significant			
	Female	Reference			
Age	30 and under	-1.5561	0.2482	<.0001	0.211
	31-45	-1.4093	0.2438	<.0001	0.244
	46-55	-1.0855	0.2475	<.0001	0.338
	56-65	Not significant			
	66 and over	Reference			
Education	Less than HS	-0.7021	0.1264	<.0001	0.496
	High School	-0.6985	0.1125	<.0001	0.497
	Some College	-0.5654	0.1108	<.0001	0.568
	Bachelor's Degree	-0.3522	0.1158	0.0024	0.703
	Graduate Degree	Reference			
Region	Northeast	Not significant			
	South	-0.2859	0.0671	<.0001	0.751
	West	Not significant			
	Midwest	Reference			
Housing Tenure	Own	0.6381	0.0603	<.0001	1.893
	Rent	Reference			
Urbanicity	CBSA, Central City	-0.4126	0.1017	<.0001	0.662
	CBSA, Not in CC, Urban	-0.3265	0.0987	0.0009	0.721
	CBSA, Not in CC, Rural	Not significant			
	Non-CBSA, Urban	Not significant			
	Non-CBSA, Rural	Reference			
Income	Missing	-0.2669	0.0896	0.0029	0.766
	Less than $20,000	-0.3792	0.1010	0.0002	0.684
	$20,000-$39,999	Not significant			
	$40,000-$74,999	Not significant			
	$75,000 or more	Reference			

Table 2: Cooperation Model Results

Variable	Category	Parameter	Std. Error	Pr>ChiSq.	Odds Ratio
Gender	Male	-0.1964	0.0539	0.0003	0.822
	Female	Reference			
Age	30 and under	Not significant			
	31-45	-0.4590	0.1468	0.0018	0.632
	46-55	Not significant			
	56-65	Not significant			
	66 and over	Reference			
Education	Less than HS	Not significant			
	High School	-0.4292	0.1036	<.0001	0.651
	Some College	Not significant			
	Bachelor's Degree	Not significant			
	Graduate Degree	Reference			
Region	Northeast	Not significant			
	South	Not significant			
	West	Not significant			
	Midwest	Reference			
Housing Tenure	Own	Not significant			
	Rent	Reference			
Urbanicity	CBSA, Central City	-0.3750	0.1023	0.0002	0.687
	CBSA, Not in CC, Urban	-0.3598	0.0988	0.0003	0.698
	CBSA, Not in CC, Rural	Not significant			
	Non-CBSA, Urban	Not significant			
	Non-CBSA, Rural	Reference			
Income	Missing	-0.7968	0.0831	<.0001	0.451
	Less than $20,000	Not significant			
	$20,000-$39,999	Not significant			
	$40,000-$74,999	Not significant			
	$75,000 or more	Reference			

Table 3: Overall Response Model Results

Variable	Category	Parameter	Std. Error	Pr>ChiSq.	Odds Ratio
Gender	Male	-0.2098	0.0434	<.0001	0.811
	Female	Reference			
Age	30 and under	-0.8323	0.1412	<.0001	0.435
	31-45	-0.7538	0.1348	<.0001	0.471
	46-55	-0.4872	0.1347	0.0003	0.614
	56-65	Not significant			
	66 and over	Reference			
Education	Less than HS	-0.6036	0.0982	<.0001	0.547
	High School	-0.6985	0.0834	<.0001	0.497
	Some College	-0.4704	0.0813	<.0001	0.625
	Bachelor's Degree	-0.2307	0.0837	0.0058	0.794
	Graduate Degree	Reference			
Hours Worked	NILF/Unemployed	Not significant			
	Less than 35 hrs/wk	0.2528	0.0962	0.0086	1.290
	35-49 hrs/wk	Not significant			
	50 or more hrs/wk	Not significant			
	Hours vary	Reference			
Region	Northeast	-0.2274	0.0659	0.0006	0.797
	South	-0.1981	0.0548	0.0003	0.820
	West	-0.1702	0.0636	0.0075	0.843
	Midwest	Reference			
Housing Tenure	Own	0.3275	0.0533	<.0001	1.387
	Rent	Reference			
Urbanicity	CBSA, Central City	-0.3667	0.0802	<.0001	0.693
	CBSA, Not in CC, Urban	-0.2863	0.0766	0.0002	0.751
	CBSA, Not in CC, Rural	Not significant			
	Non-CBSA, Urban	Not significant			
	Non-CBSA, Rural	Reference			
Income	Missing	-0.6939	0.0711	<.0001	0.500
	Less than $20,000	Not significant			
	$20,000-$39,999	Not significant			
	$40,000-$74,999	Not significant			
	$75,000 or more	Reference			
Presence of Children	With Children	0.1398	0.0508	0.0059	1.150
	No Children, 1 Adult	Not significant			
	No Children, 2 or More Adults	Reference			