

SUPPORTING STATEMENT PART B

B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS

1. Respondent Universe

The ATUS sample is drawn from the Current Population Survey (CPS), so the ATUS universe is the same as the CPS universe. The universe for the CPS is composed of the approximately 118 million occupied households in the U.S. and the civilian, noninstitutional population residing in those households. From this universe, the CPS sample includes approximately 60,000 eligible households every month. About one-eighth (or about 7,500) of these retire permanently from the CPS sample each month after their eighth CPS interview attempt. Households that complete their eighth-month interview are eligible for selection for ATUS. About 2,060 of the households in this group that complete their eighth CPS interview will be selected for the ATUS sample each month.¹ On average, about 100 households will be identified as ineligible; designated respondents may have moved or died or the household may be ineligible for another reason. Based on the average response rate over 2003-15, a response rate of about 49.8 percent is expected over an 8-week fielding period. Thus, about 980 interviews will be completed each month (1960 eligible respondents x .498). In 2015, about 433 or 36 interviews per month were then thrown out of the estimation process because they did not meet ATUS minimal data quality standards.²

The ATUS has a stratified, three-stage sample. In the first stage of selection, the CPS oversample in the less populous States is reduced. The CPS is designed to produce reliable estimates at the State and national level. The ATUS does not have a State reliability requirement. Because of the CPS State reliability requirement, the less populous States are allocated a larger proportion of the national CPS sample than they would get with only a national reliability requirement. In order to improve the efficiency of the national estimates from the ATUS, the CPS sample is subsampled to obtain the ATUS sample. The sample that remains after the subsampling is distributed across the States is approximately equal to the proportion of the national population they represent.

¹ In 2003, the first year of full production, the ATUS sample was 35 percent higher than in later years. The original target was to complete 2,000 interviews per month. The monthly sample was reduced beginning in December 2003 in order to bring survey costs in line with the survey budget. The original annual sample was drawn to meet the target goal assuming a 70% response rate. The goal was twice the minimum 12,000 interviews/year (1,000/month) originally identified by Robison (1999), in "Sampling and Reporting in Time-Use Surveys," as the number required to contrast time-use estimates for major subpopulations of interest. Robison recommended adding an additional 12,000 interviews to enable more subpopulation comparisons. His assumptions used time-use distributions for various subpopulations from a 1975 University of Michigan time-use survey as well as associated parameters that enabled the calculation of standard errors and confidence intervals under different assumptions. These numbers and parameters were published in Juster and Stafford (1985).

² Interviews that are thrown out of the estimation process for not meeting minimum data quality standards include those that 1) have fewer than 5 activities, or 2) have more than 180 minutes of "don't know" or "refused" activities, or 3) both.

In the second stage of selection, households are stratified based on the following characteristics: race/ethnicity of householder, presence and age of household children, and the number of adults in adult-only households. Sampling rates vary within each stratum. Eligible households with a Hispanic or non-Hispanic black householder are oversampled to improve the reliability of time-use data for these demographic groups. To ensure adequate measures of childcare, households with children are also oversampled. To compensate for this, households without children are undersampled.

In the third stage of selection, an eligible person from each household selected in the second stage is selected as the designated person (respondent) for the ATUS. An eligible person is a civilian household member at least 15 years of age. All eligible persons within a sample household have the same probability of selection.

The sample persons are then randomly assigned a designated reference day (a day of the week for which they will be reporting) and an initial interview week code (the week the case is introduced). In order to ensure accurate measures of time spent on weekdays and weekend days, the sample is split evenly between weekdays and weekend days. Ten percent of the sample is allocated to each weekday and 25 percent of the sample is allocated to each weekend day.

The following table shows the approximate number of households sampled annually from each stratum.

Table 2 Estimated annual sample size by ATUS sampling strata, 2004 and later

Household type	Race/ethnicity of household reference person in CPS			Total
	Hispanic	Non-Hispanic, black	Non-Hispanic, nonblack	
With at least one child under 6	1,200	600	3,400	5,200
With at least one child between 6 and 17	1,200	900	4,900	7,000
Single adult, no children under 18	700	1,600	4,300	6,600
Two or more adults, no children under 18	1,200	1,400	5,000	7,600
Total	4,300	4,500	17,600	26,400

Estimation includes a series of adjustments to account for the stages of sample selection, a non-response adjustment, and a benchmarking procedure which will ensure that certain quarterly population counts from the ATUS sample agree with corresponding counts from the CPS.

The initial weight for each ATUS sample case is the CPS weight after the first-stage adjustment. This weight accounts primarily for the probability of selecting the

household for the CPS and for CPS non-response. This weight is then adjusted by three factors to account for: the reduction of the CPS oversample in less-populous States, the probability of selecting the household within the ATUS sampling strata, and the probability of selecting the individual person from each sample household. The non-response adjustment increases the weights of the responding sample cases to account for those who didn't respond by reference day and incentive status. Additional details on the weighting procedures are provided in the ATUS Weighting Plan (see Attachment J).

The benchmarking procedure is an iterative raking procedure containing three steps. The first step adjusts the weights of the sample cases so that weighted estimates of persons in various gender-race/ethnicity categories from the ATUS agree with similar population counts from the CPS. The second step of the benchmarking procedure adjusts the weights of the sample cases so that estimates from the ATUS match composite estimates from the CPS for household composition and educational attainment by gender. The third step adjusts the weights so that weighted estimates by age category and gender agree with CPS population counts. In all three steps, weights are adjusted separately for weekdays and weekend days so that population estimates agree with CPS for both day-of-week categories.

The probability that an individual participates in an activity on a given day varies across activities. For example, nearly everyone reports sleeping on the diary day, while few people report educational activities. A balanced repeated replication variance estimator is used to calculate standard errors and coefficients of variation for selected estimates. Table 3 shows the coefficients of variation (CV) of ATUS quarterly and annual average (2015) hours measures for activity categories that were published in the release of ATUS estimates.

Table 3 Quarterly and annual average CVs on average hours estimates, 2015

Activity	Estimated average CV, Quarterly estimates, 2015	CV Annual estimates, 2015
Personal care, including sleeping	0.006	0.003
Eating and drinking	0.018	0.01
Household activities	0.029	0.014
Purchasing goods and services	0.04	0.021
Caring for and helping household members	0.049	0.024
Caring for and helping non-household members	0.102	0.048
Working and work-related activities	0.029	0.014
Educational activities	0.128	0.054
Organizational, civic, and religious activities	0.07	0.034
Leisure and sports	0.016	0.008
Telephone calls, mail and email	0.081	0.04
Other activities, n.e.c.	0.1	0.05

2. Description of Procedures

A. Estimation Procedures

Four types of estimates are used to produce published ATUS tables: average hours per day, participation rates, number of participants, and average hours per day of participants.

Average Hours per Day: The average number of hours spent per day engaging in activity j for a given population, \bar{T}_j , is given by

$$\bar{T}_j = \frac{\sum_i fwt_i T_{ij}}{\sum_i fwt_i}$$

where T_{ij} is the amount of time spent in activity j by respondent i , and fwt_i is the final weight for respondent i .

Participation Rates: The percentage of the population engaging in activity j on an average day, P_j , is computed using

$$P_j = \frac{\sum_i fwt_i I_{ij}}{\sum_i fwt_i}$$

where P_j is the percentage of people who engaged in activity j in a given day, and I_{ij} is an indicator that equals 1 if the respondent i engaged in activity j during the reference day and 0 otherwise.

In this type of estimate, P_j does not represent the proportion of people who participate in activity j over periods longer than a day.

Number of Participants: The number of persons engaging in activity j during an average day, Num_j , is given by

$$Num_j = \sum_i \frac{fwt_i I_{ij}}{D}$$

where Num_j is the number of persons participating in activity j during an average day,

I_{ij} is an indicator that equals 1 if respondent i participated in activity j during the reference day and 0 otherwise, and

D is the number of days in the estimation period (365 for annual averages for non-leap years, for example).

Average Hours per Day of Participants: The average number of hours spent per day engaged in activity j by participants, \bar{T}_j^P , is given by

$$\bar{T}_j^P = \frac{\sum_i fwt_i I_{ij} T_{ij}}{\sum_i fwt_i I_{ij}}$$

where T_{ij} is the amount of time spent in activity j by respondent i , and

fwt_i is the final weight for respondent i , I_{ij} is an indicator that equals 1 if respondent i participated in activity j during the reference day and 0 otherwise.

B. Data Collection

The ATUS interview is a combination of structured questions and conversational interviewing. For the household roster update, employment status questions, and CPS updates, Census Bureau interviewers read the question on the screen and enter the appropriate response. For the time-use “diary” and subsequent summary questions on childcare, paid work, volunteering, and eldercare, the interviewer interviews the respondent more flexibly, filling in the diary grid as questions are answered. The data collection instrument includes an edit check that ensures that all cells are filled before the interviewer exits the diary. Extensive interviewer training has been provided in how to do conversational interviewing—including

when to selectively probe for adequate information to code activities. Refresher training is conducted at least annually. Interviews are routinely monitored by supervisors, coaches, and BLS sponsors to evaluate conversational interviewing performance. The coding task helps to ensure that interviewers understand the level of detail needed in activity reports for accurate coding; all interviewers are also coders, though interviewers do not code their own work. A coding verification and adjudication process is in place. Verification continues to be done at 100 percent to ensure high and consistent data quality.

3. Methods to Maximize Response

A number of efforts have been undertaken to maximize ATUS survey response rates.

- A. *Field Test*. The 2001 field test examined the effectiveness of incentives, sending advance materials by priority mail, doubling the number of eligible interviewing days by using a day-of-week substitution methodology, calling in advance to set interview appointments, “recycling” cases for field visits, and extending the field period from 4 to up to 8 weeks. (See Attachment B.)
1. *Use of Incentives and recycling cases to the field*. As discussed in Part A, section 9, testing showed that incentives significantly increased response rates. “Recycling” cases to the field—that is, turning nonresponse cases over to interviewers to conduct face-to-face interviews in the respondent’s home—was also effective in maximizing response rates, particularly for no-telephone-number households. However, incentives to all respondents and recycling were both cost prohibitive.
 2. *Appointment setting*. Calling in advance to set an appointment (“proactive appointment setting”) did not improve response, and completed interviews using that strategy required 70 percent more contact attempts than other completed interviews. As a result, advance appointment setting was rejected.
 3. *Day-of-week substitution*. Allowing day-of-week substitution increased response rates by about 4 percentage points over 8 weeks; however, this practice led to a disproportionately high number of completed interviews on Wednesdays and a disproportionately low number on Fridays. To maintain integrity in the day-of-week distribution of the sample, substitution was also rejected.
 4. *Use of priority mail*. Consistent with survey methods literature, priority mail appears to have increased response rates in the ATUS field test—by over 10 percentage points. It is relatively low cost to implement (\$6.45 per mailing in 2016) and is currently used for sending advance materials.
 5. *Fielding period*. The optimal field period length varies depending on incentive use. Without an incentive, the field test showed that an 8-week fielding period was required to approach 70 percent (69 percent was achieved

American Time Use Survey (ATUS)

1220-0175

October 2016

in the field test). As a result, this 8-week fielding period was adopted for full production. To even out workload and measure time use across days of the month, one quarter of the monthly sample is introduced each week for 4 weeks. Active cases are called up to 7 times per day on one eligible day each week for 8 weeks.

- B. *Incentive expansions.* In addition to sending \$40 incentives to individuals in households for which the Census Bureau does not have a phone number, two OMB-approved incentive expansions were implemented in recent years. As of 2013, incentives are sent to DPs in no-telephone-number households as well as individuals for whom the Census Bureau assigned call outcome codes of: *108 Number not in service*; *109 Number changed, no new number given*; *124 Number could not be completed as dialed*; and *127 Temporarily not in service* after the first week of collection. (See Attachment C.) The use of incentives has helped to boost response among difficult-to-reach populations. Individuals who are sent incentives are more likely to be black, of Hispanic or Latino ethnicity, to have less education, and to have lower household incomes than members of households that provide phone numbers.
- C. *Toll-free number provided to DPs.* To maximize response, a toll-free number is provided to all eligible respondents in the advance materials. They can use the number to call in and set an appointment for an interview or, if they call on their interview day, to complete the interview.
- D. *Advance materials revised.* In 2005, an examination of the ATUS advance materials was undertaken and the advance materials were subsequently revised. The advance materials were reviewed and updated again in 2012-13. The advance letters were revised to include information commonly asked by respondents during their first contact with interviewers. The ATUS brochure was updated and redesigned to appeal to more respondents. The debit card and instruction sheet also were redesigned to appear more prominently in the advanced mailer envelope. These materials were modified based on feedback received from expert reviewers and focus groups of ATUS interviewers who examined existing materials. (See Attachments E, I, and J).
- E. *Respondent Web site.* BLS developed a Web site to address common respondent questions about the survey. Its web address is included in the advance letters (<http://www.bls.gov/respondents/tus/home.htm>).
- F. *Fax letters.* BLS worked with Census to develop "we've been trying to reach you letters" to fax to telephone numbers that reach fax machines. Like an answering machine message, the fax letters ask the sampled person to call the Census Bureau and complete an interview.
- G. *Call boost.* Interviewers adjust WebCATI settings such that a call is likely to occur around the same time as final contact was made in the last CPS interview. This practice is supported by research that showed contact was more likely if ATUS calls were made around the same time as previous CPS contact, even if the ATUS DP is a different person than the CPS MIS-8 respondent.

- H. *Interview Operations Analysis.* In 2004, telephone call center operations were examined to determine if measures could be taken to increase response rates, and three basic operations were changed. First, the ATUS staff learned that while many surveys set calling goals for interviewers, the call center management was not providing ATUS interviewers with daily or weekly goals. Beginning in the summer of 2004, the telephone center management set daily goals for ATUS interviewers, providing concrete guidelines for how many completed calls are desired. Although the interviewers do not always meet their goals, these goals assist the telephone center management to measure daily progress and to motivate the interviewers. Second, it was discovered that because of the way call blocks (times) were scheduled, many calls were being made between about 4:30 pm and 5:00 pm, before many people were home from work. Methods for calling were changed so that more calls would be made after 5:30 pm, when people who work regular 9-5 hours would be more likely to be home. Finally, the Census Bureau conducted more research into invalid phone numbers in an attempt to find valid phone numbers for the contact person.
- I. *Interviewer job aids.* Interviewers have job aids—answers to frequently asked questions—designed to help answer questions about the survey and to assist them in gaining respondents' cooperation to participate.
- J. *Interviewer incentives.* An interviewer incentive study was considered but subsequently rejected as the reality of implementing interviewer incentives was determined to be cost prohibitive.
- K. *Newsletters.* In cooperation with Census, BLS periodically produces newsletters that are designed to motivate and inform interviewers.
- L. *Interviewer training.* BLS and Census have conducted workshops for interviewers on techniques to gain cooperation from respondents, and much of the material developed for this training was incorporated into other interviewer training courses. Interviewer operations also have been scrutinized and revised to increase the probability of completed interviews, such as redesigning the call blocks to add more call attempts during evening hours.
- M. *Studies to understand nonresponse and possible nonresponse bias.* In addition to the efforts listed above, a number of studies have been done to understand nonresponse in the ATUS. More detail about these studies appears in Section 4, Tests and Research.

4. Testing of Procedures

Before the ATUS went into full production, extensive testing was done on the operations methodologies, question wording and interpretation, and activity coding. All questions added to the survey over the years also were subject to extensive testing before their implementation.

A. Completed research

1. *Operations Field Test*. The ATUS presents special operational challenges because a designated person—rather than any household member—must be contacted on a specific day of the week. The field test was designed to examine methods to maximize respondent contact and response. The 2001 operations field test is mentioned throughout this clearance package and is described in more detail in Attachment B.

2. *Cognitive testing*

a. **Diary**

None of the completed cognitive tests focused specifically on the time diary, although the ATUS introduction, instructions, roster update, time diary and associated contextual information were administered as part of all tests. As a result, respondents' reactions to each of these survey elements were used to modify and improve the survey. Modifications based on respondent reactions include:

Time diary instructions

Time diary instructions were shortened so that they take approximately one minute to administer. Two major modifications were made to the original instructions: the original instructions did not specify that respondents needed to estimate the duration of each activity. As a result, respondents often "laundry-listed" activities without attributing times to each activity. Language informing respondents to estimate activity duration was added to the time diary instructions. As a result, fewer respondents have required prompting to provide time estimates. The original instructions included examples of how to report activities. Research showed that these examples were not helpful to respondents because they failed to match respondents' daily circumstances (Stinson, 2000). Dropping examples from the time diary instructions shortened the time it took to administer them, and of the nearly 100 people who participated, fewer than 10 respondents requested examples that would specify the level of detail needed in the time diary.

"Who was with you?"

Stinson (2000) and Schwartz & Fricker (2000) found that the question, "Who was with you?" was open to multiple interpretations. Some respondents interpreted the question as meaning, "Who was near you?" whereas others understood it to mean, "Who participated in the activity with you?" In order to make the probe clearer, ATUS interviewers ask, "Who was in the room with

you?” when respondents are at their own or someone else’s home. They ask, “Who accompanied you?” for activities that occur in other locations. Respondents are not asked “Who was in the room with you?” when they report sleeping, grooming, personal activities, or being at work. In 2008, the questions “Who was with you?” and “Who accompanied you?” were cognitively tested for times when respondents reported working or doing work-related activities. None of the respondents had difficulty remembering who was with them while they were working, although some respondents did not provide the level of detail that was desired. To ensure an appropriate level of detail is collected, respondents who say they were with “co-workers” are asked the follow-up question, “By co-workers, did you mean you were with your manager/supervisor, people whom you supervise, or other co-workers?”

b. Childcare

Focus groups and two rounds of cognitive testing were conducted to refine the wording of the childcare summary question (Schwartz & Fricker, 2000; Schwartz, 2001). Based on the findings from those studies, reports of care for household children are restricted to times during which at least one child under the age of 13 was awake. The phrase “*in your care*” was selected to convey that the parent or care provider was responsible for and mindful of at least one child 12 years old or younger. For more details, please see the summary of cognitive lab #2106 that was provided to OMB in July 2001.

c. Paid work

Stinson (2000) and Schwartz, Lynn & Gortman (2001) conducted three rounds of cognitive testing of the paid work summary questions. The major findings were that respondents interpreted both concepts, *activities done for one’s job or business* and *activities done for pay*, more broadly than researchers had intended. Based on respondents’ reports, *activities done for one’s job or business* can include networking or relationship-building activities and *activities done for pay* can include any income-generating activity that is not one’s main or second job.

d. Eldercare Questions

In January 2011, questions on eldercare were added to the ATUS. These questions were cognitively tested on both caregivers of the elderly and the general public, and the results of these tests were used to refine the wording of the eldercare summary questions. The questionnaire was tested for clarity, comprehension, length, potential sensitivity, and the flow through the instrument. Respondents were asked if they have provided care or assistance to an adult who needed care because of a condition related to aging. The phrase “*condition related to aging*” was selected because the focus group results and research showed disagreement on a specific age that eldercare begins. Cognitive testing of the phrase and the questions showed the wording was effective in identifying individuals who had provided care to the elderly. (See Attachment N.)

3. *Coding Tests.* The ATUS coding lexicon was developed for and is unique to the ATUS. While originally based on the system used in the Australian Time-Use Survey, the system was modified a great deal to enable more detailed and flexible analysis of time-use data. Modifications were driven by results of four coding tests and by issues brought up in production. The first 3 tests were conducted with Census Bureau coders and the fourth with Westat coders. The tests examined the intuitiveness of the coding system, accuracy rates by activity tier, inter-coder variability, and coding software usability. A systems test of the coding verification and adjudication process was also completed in October 2001.

4. *Software tests.* The ATUS data collection instrument is programmed in modules or blocks using Blaise software. Each block was extensively tested at Census and BLS prior to full production. Testing scenarios were repeated with each version of the instrument prior to production, and additional testing scenarios are run any time a change is made to the instrument to ensure that all modifications are correct and that there are no unintended consequences. "Audit trails" capturing every key stroke are used to investigate problems. Instruments are also tested by Census Bureau interviewers prior to being used.

5. *Advance diary test.* Early in ATUS development, survey methodologists recommended sending diaries with the ATUS advance materials to facilitate recall and improve data quality. There was some concern among the survey sponsors about sending diaries in advance without testing effects on response.

BLS awarded a contract to the National Opinion Research Center (NORC) to conduct a split-panel test of advance diaries in April 2002. Half of the respondents in this test (n =225) received an advance diary and then completed a telephone interview that used conversational interviewing to elicit the details needed for coding. The other respondents (n =225) received the same advance materials with the exception of the diary and engaged in the standard time-use interview. NORC found that sending an advance diary increased burden, and did not improve data quality or response. The NORC final report was sent to OMB in December 2003.

After receiving the NORC test results, the BLS Office of Survey Methods Research further analyzed the data using multivariate analyses. This analysis confirmed NORC's results. As a result, no diary was added to the ATUS advance materials.

6. *Simultaneous activities.* Secondary or simultaneous activities are considered one of the significant dimensions of an activity that should be captured in a time diary (Harvey, 2001). Early research at BLS as well as experience by Statistics Canada indicated that the systematic collection of secondary activities could be problematic in a telephone survey. While a paper diary form simply needs to include a column for secondary activities in order for respondents to know that they should record them, in a telephone survey, interviewers must probe, "*Were you doing anything else?*" for each activity in order to collect information in a systematic and unbiased way. Probing for secondary activities can quickly become burdensome and introduces the risk of fatiguing the respondent early in the interview. Additionally,

Stinson (2000) found that respondents could not attribute times to secondary activities, which would weaken their analytical relevance. Nevertheless, research participants, members of advisory councils, and survey methodologists have all recommended collecting simultaneous activities.

In 2003, BLS solicited proposals from NORC to look at the systematic collection of simultaneous activities. The study was necessarily complex and costly. BLS decided to delay cognitive work on this subject until some empirical data on simultaneous activities were available from full production.

ATUS interviewers ask respondents to report the main activities they did on the diary day. From 2003-12, when interviewers voluntarily provided info about simultaneous activities the interviewers recorded but did not code this information. Activity codes were assigned to some of these data for a 2011 research study that revealed ATUS respondents' infrequent reports of secondary activities accounted for much less time compared with traditionally-collected secondary activity reports (Drago, 2011). The study went on to conclude that the ATUS data on simultaneous activities were of low quality and limited value; because of this, in early 2012, Census interviewers stopped recording voluntary reports of simultaneous activities in the ATUS.

7. Advance Materials Analysis. In 2004, two studies were undertaken to re-examine the ATUS advance materials. An expert review of the materials and focus groups with ATUS interviewers were conducted to determine how the advance materials might be re-designed to better influence designated persons to participate. Findings from both studies indicated the letter should be shorter and the brochure should have a more appealing design, including switching from a dichromatic to a full color scheme. In addition, the focus groups and expert reviewers recommended revising the brochure to address respondents' questions. In 2005, extensive revisions were made to the advance materials based on these studies. (See Attachment O.) The advanced materials were extensively reviewed, modified, and updated again in 2012-13. Changes were made based on feedback from expert reviewers and focus groups of ATUS interviewers.

8. Incentive experiment. In line with terms of clearance from the 2003 OMB package, the feasibility of an incentive experiment conducted in a production environment was considered. A BLS and Census Bureau interagency team discussed the development of an experiment, with the intention of conducting it in fiscal year 2005. Planning and assessment meetings determined that the incentive experiment was not a viable option for increasing response rates due to the costs associated with providing incentives to all ATUS participants.

9. Item nonresponse. BLS investigated the incidence of missing and imputed ATUS data to assess the quality of ATUS variables. Item nonresponse was found to be quite low in the ATUS, with most variables having an item nonresponse of well under 2 percent. The two variables describing weekly and hourly earnings had

higher incidences of nonresponse compared to other variables (see chapter 6 of the ATUS User's Guide at <http://www.bls.gov/tus/atususersguide.pdf>).

10. *Cell phone response analysis.* Meekins and Denton (2012) used the ATUS to examine the impact of calling cell phone numbers on nonresponse and measurement error. They found that cell phone respondents have higher noncontact rates, but have refusal rates that are similar to landline respondents. They also note that there is no significant difference in the measurement error rates or in estimates of time use for both groups. (See Attachment E.)

11. *Call block research.* Meekins (2013) looked at call patterns to determine whether greater efficiencies could be attained without biasing ATUS data. Using ATUS call history data from 2006 to 2007, he found that a small number of ATUS sample units receive a disproportionately large amount of effort. His results also showed that dialing around the same time as a previous contact is a positive predictor of subsequent contact while refusing to provide income information on the CPS is a negative predictor and calling efficacy is greater in the later hours of the day. The study concluded with several recommendations for optimizing the efficiency of calls to sample units. (See Attachment P.)

12. *Behavior coding.* Behavior coding is a technique that has been successfully utilized with event history calendar data collection (Belli, 2004) to understand how interviewers ask questions and provide clarification and feedback to respondents, how respondents interpret questions and recall answers, and how interviewers and respondents interact during the survey task. ATUS interviewers are trained in conversational interviewing techniques, which allow for interventions with a respondent to help him or her stay on track when remembering the day's activities, and activity sequences and timing. Research from this study evaluated how conversational interviewing and specific recall techniques are used by interviewers and their effect on data quality, which will help aid in instrument development and interviewer interventions. (See Attachment Q.)

13. *Recall period research.* BLS worked with outside researchers to examine whether an extended recall period affects data quality. As a whole, the research project examined various data quality measures—for example, the number of activities per day, the percent of “good” or “bad” time diaries—by the length of the recall period. Results from the analysis of ATUS data revealed no major differences in data quality that could be attributed solely to the length of the recall period. However, results from investigating other time diary surveys \ demonstrated some declines in data quality with longer recall periods; for example, respondents who said they completed the diary more than 24 hours after the diary day had 14-25% fewer activities in their diaries as respondents who said they completed their diaries while they went about their days. With the indication that data quality might be negatively affected and the increased costs and managerial challenges associated with developing and managing a collection system that uses two recall periods, BLS decided not to pursue additional research on an extended recall period.

14. *Web collection of ATUS diary.* BLS consulted with Westat to explore the feasibility of using a mixed-mode design that includes the collection of ATUS data via a Web instrument. A move to a mixed-mode design could potentially help ATUS improve response and be prepared for the survey climate of the future. The project included a literature review of web and mixed-mode data collection, provided recommendations on the design of web data collection for the ATUS, including respondent allocation and contact strategies and question design considerations for a web instrument. The project also included a discussion of comparability issues between web and telephone data collection with methods to evaluate the proposed design including errors of nonobservation (e.g., coverage and nonresponse error) and errors of observation (e.g., measurement error). Westat also provided a preliminary mockup of the recommended diary design. (See Attachment R.) While ATUS is not pursuing web collection at this time, the project will help inform any potential issues of introducing web data collection should ATUS do so in the future.

15. *Research done on ATUS response and nonresponse.* Numerous studies have been done to understand ATUS survey response and nonresponse. BLS, the Census Bureau, and researchers who are not affiliated with these agencies all have been active in this area.

a. Census Bureau Response Rate Investigation. A team at the Census Bureau compared response rates achieved in the beginning of 2003 with higher rates achieved in 2002, just before ATUS full production began. The team tested several hypotheses in an attempt to determine why response declined at that time. The team examined whether there were changes in the number or timing of call attempts, and whether the hiring of new interviewers just before full production or problems with the call scheduling software might have affected response. While they found some spikes in times of day that people refuse, they did not find a strong pattern for day of week or time of day effects in refusal rates. They also found that there was no relationship between interviewers' ATUS refusal rates and their years of experience interviewing. In a multivariate analysis, the team found a correlation between a refusal to provide income data in CPS and a refusal to participate in ATUS. This information could be valuable for predicting nonresponse and/or targeting refusal conversion efforts. (See Attachment K.)

b. Response Analysis Survey. In 2004, qualitative research was completed to look at reasons for nonresponse in ATUS. In January 2004, the BLS developed and the Census Bureau conducted a Response Analysis Survey (RAS). Census Bureau interviewers attempted to contact a sample of both respondents and non-respondents to the ATUS to learn more about persons' propensities to respond or not to the ATUS, and to better understand to which features of the survey response propensity might be correlated. The study focused on refusals rather than noncontacts, as the former are the main contributor to ATUS non-response. It was restricted to English-speaking adults selected for the ATUS. The primary reason that RAS respondents mentioned for not participating in the ATUS was that they were tired from responding to the CPS. The RAS also

included questions about whether respondents read the advance materials, visited the web site, or sent e-mails asking for information, as well as their impressions of Census Bureau interviewers. Based on the responses to the RAS, the BLS examined how to best alter survey operations to increase designated persons' propensities to respond. Advance materials were revised to explain more clearly the reasons why some CPS respondents were "re-selected" for the ATUS, and the ATUS brochure was redesigned to increase the proportion who read it, and to feature the web site address more prominently. The RAS report is included with this supporting statement (available on the Internet at <http://www.bls.gov/ore/pdf/st040140.pdf>).

c. Alternative contact strategies. Using simulated data, Stewart (2002) examined the effect of using different contact strategies in a telephone survey. He found that allowing for day-of-week substitution resulted in a systematic bias, and that data collected would overstate the amount of time spent away from home. By contrast, a designated-day approach resulted in little bias.

d. Analysis of returned mail. Census Bureau staff conducted an analysis of returned advance mailings and postcards to assess how effective their address review and correction process was, what the impact on response rates would be if addresses identified as movers were reassigned as "not eligibles," and how the mail return rates differed between incentive and non-incentive cases. The authors considered reassigning the 06 mover codes, the 08 address correction provided codes, and other codes. The research concluded that converting all returned mail cases currently coded as eligible to not eligible would only improve the overall response rate by a maximum of 1.22%, about half of which would be due to the 06 and 08 codes. It was also discovered that twice as many incentive cases had advance mailings returned than non-incentive cases, and those cases that had the advance mailings returned were three times less likely to complete the ATUS interview. Incentive cases are a special concern because respondents must contact the call center to complete the interview, and this contact information is provided in the advance letter. In order to increase incentive case response rates, Census Bureau staff now researches addresses for all incentive cases that had mail returned. (See Attachment M.)

e. Substitution of DPs and of diary days. In 2012-13, BLS contracted with Westat to provide guidance on some methodological changes that could be made to the ATUS with the intent of increasing response rates. The two options considered were allowing for a substitution of diary days and a substitution of designated persons (DPs) if the DP is not available on the assigned diary day. Westat found evidence that some day-of-week substitutions might successfully raise response rates without inducing bias; for example, they found that Monday – Thursday diaries are relatively interchangeable. Westat designed an experiment that could test these theories. Westat advised against allowing a substitution of DPs in the ATUS design because of concerns that doing so may bias the data. (See Attachment F.)

f. Nonresponse bias analyses. BLS, the Census Bureau, and outside researchers have completed a number of nonresponse bias analyses over the years. In 2005, O’Neil and Dixon conducted an in-depth analysis to examine patterns of ATUS non-response using CPS data. This analysis included breaking out non-response by a variety of demographic characteristics, using logistic analysis to determine variables related to non-response, and building a propensity score model to examine differences in time-use patterns and to assess the extent of non-response bias. Findings showed race and age to be strong predictors of ATUS refusals and noncontacts. The study also showed that estimates of refusal and noncontact bias were small relative to the total time spent in the activities. A follow-up to this analysis (Dixon, 2006) found no nonresponse biases in the time-use estimates, probability of use of time categories, or the relationship between the categories. The study further concluded that any potential biases identified were small.

The ATUS survey methodology files are available to the public, enabling outside researchers to examine survey methods issues. Abraham et al. (2006) found that people weakly integrated into their communities were less likely to respond to ATUS, mostly because they were less likely to be contacted. They also found little support for their hypothesis that busy people are less likely to respond to the ATUS. The authors compared aggregate time use estimates using the ATUS base weights without adjustment for nonresponse, using the ATUS final weights with a nonresponse adjustment, and using weights that incorporated the authors’ nonresponse adjustment based on a propensity model. They found the three sets of estimates to be similar.

Letourneau and Zbikowski analyzed nonresponse in the ATUS using 2006 data (2008). Some results from this study were consistent with previous non-response bias studies, such as lower response rates for those living in urban areas and higher refusal rates for those missing the CPS income variable. However, this study contradicted previous studies in several areas. Contrary to previous studies, this Census study did not find lower response rates for the unemployed or those not in the labor force. It also did find lower contact rates for people who work longer hours, and for blacks and Hispanics.

A 2009 paper (Abraham, Helms, and Presser, 2009) found that ATUS respondents were more likely to be volunteers than the general population, and that therefore the ATUS estimate of volunteer hours is biased upward. The authors estimated the associations between respondent characteristics and volunteer hours, and found them to be similar to those from the CPS Volunteer Supplement.

Fricker and Tourangeau examined characteristics that affect nonresponse using 2003 data (2010). Many of their findings were consistent with earlier studies regarding age, race, income, and respondent busyness on response rates. They found higher nonresponse for those who skipped the CPS family income questions, had been a CPS nonrespondent, or were not the respondent in the

last CPS interview. The authors also found that removing cases with a high nonresponse propensity from the sample produced small but significant changes in the overall time use estimates.

Dixon and Meekins focused on nonresponse bias and measurement error in the ATUS (2012). Using a propensity score model to examine differences in time use patterns and to assess the extent of nonresponse bias, the authors found the estimates of bias were very small from all sources.

Dixon used 2012 data to examine nonresponse using propensity models for overall nonresponse as well as its components, refusal and noncontact. He also explored the possibility that nonresponse may be biasing the estimates due to the amount of zeroes reported by comparing the proportion of zeroes between the groups. The author found no nonresponse bias, but did find the level of potential bias differed by activity. The differences between the reported zeroes from the survey and the estimated zeroes for nonresponse were very small, suggesting that reasons for doing the activity were likely not related to the reasons for nonresponse.

See Attachment L for a table containing information and hyperlinks for most of the research cited above.

B. In progress and planned research

1. *Interviewer/coder debriefings.* Many interviewer debriefings have been conducted since full production began in 2003. These debriefings have illuminated procedural difficulties and identified questions that interviewers feel pose problems for respondents. They also assist in clarifying interviewer questions and improving future training.

2. *Nonresponse bias research.* Earp and Edgar are currently assessing the potential for nonresponse bias in the ATUS using 2014 data by comparing the characteristics of ATUS respondents and nonrespondents using a regression tree model using demographic variables from the CPS. To determine the potential for nonresponse bias, they compare the CPS employment rate of ATUS respondents and nonrespondents, since employment status is expected to be related to time use. Preliminary results show the employment rate did not vary significantly between ATUS respondents and nonrespondents in the overall sample, indicating little to no nonresponse bias for ATUS estimates correlated with CPS employment status.

3. *Responsive design research.* The Center for Survey Measurement (CSM) at Census is investigating strategies to improve the respondent contact rate as a means of improving the survey response rate. The current non-contact rate is equal to the refusal rate but may be easier to alter through improved call time strategies and phone number research methods. CSM is investigating the use of call time data

from the CPS as well as methods for targeting population groups within the sample who have proven to be more difficult to contact.

5. Contact Persons

The following individuals may be consulted concerning the statistical data collection and analysis operation:

Statistical Design

Antoinette Lubich
Demographic Statistical Methods Division
U.S. Census Bureau

Data Collection

Beth Capps
Assistant Survey Director for the American Time Use Survey
Associate Director for Demographic Programs
U.S. Census Bureau

Statistical Analysis

Rachel Krantz-Kent
Program manager
American Time Use Survey
Bureau of Labor Statistics

Attachments:

- A - ATUS Instrument Specifications
- B - Field Test Analysis
- C - Incentive Expansion OMB Memo
- D - ATUS Debit Card Mailer
- E - Cell Phone Research
- F - Substitution of Days Study
- G - Legal Authority Backing
- H - Advance Letters
- I - Advance Brochure
- J - ATUS Weighting Plan
- K - Response Rates Analysis
- L - Summary of Nonresponse Bias Studies
- M - Returned Mail Analysis
- N - Eldercare Cognitive Summary Report
- O - Advance Materials Reevaluation
- P - Call Block Research Study
- Q - Behavior Coding Research
- R - ATUS Web Collection Study
- S - Table A7 Median Hourly Earnings

Works Cited

Abraham, Katharine G., Aaron Maitland, and Suzanne M. Bianchi. (2006). "Nonresponse in the American Time Use Survey: Who Is Missing from the Data and How Much Does It Matter?" *Public Opinion Quarterly*, 70 (5): 676-703.

Abraham, Katharine G., Sara E. Helms, and Stanley Presser. "How Social Processes Distort Measurement: The Impact of Survey Nonresponse on Estimates of Volunteer Work." *American Journal of Sociology*, January 2009.

American Association for Public Opinion Research. (2008). Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 5th edition. Lenexa, Kansas: AAPOR.

Belli, R., Lee, E.H., Stafford, F.P., and Chou, C.H. (2004). "Calendar and Question-List Survey Methods: Association Between Interviewer Behaviors and Data Quality." *Journal of Official Statistics* 20: 185-218.

Bridgman, Benjamin. "Accounting for Household Production in the National Accounts: An Update, 1965-2014." *Survey of Current Business*. Volume 96, No.2/February 2016.

Drago, Robert. "Secondary Activities in the 2006 American Time Use Survey." Bureau of Labor Statistics Working Paper, February, 2011.

Dixon, John. "Nonresponse Bias for the Relationships Between Activities in the American Time Use Survey." Bureau of Labor Statistics Working Paper, 2006.

Dixon, John and Brian Meekins. "Total Survey Error in the American Time Use Survey." Bureau of Labor Statistics Working Paper, 2012.

Fricker, Scott and Lisa K. Schwartz. (August, 2001). "Reporting absences from home: Results of cognitive testing of the American Time Use Survey's missed days summary question."

Fricker, Scott and Roger Tourangeau. "Examining the Relationship Between Nonresponse Propensity and Data Quality in Two National Household Surveys." *Public Opinion Quarterly*. Volume 74, No. 5/December 2010.

Harvey, Andrew S. "Guidelines for Time Use Data Collection." *Social Indicators Research* 30 (1993): 197-228.

Herz, Diane E., Lisa K. Schwartz, and Kristina J. Shelley (2001). "Coding Activities in the American Time Use Survey."

American Time Use Survey (ATUS)
1220-0175
October 2016

Landefeld, J. Steven and Stephanie H. McCulla. "Accounting for Nonmarket Household Production Within a National Accounts Framework." *Review of Income and Wealth, Series 46, Number 3* (September 2000): 289-307.

Letourneau, Phawn M. and Andrew Zbikowski. "Nonresponse in the American Time Use Survey," published in the proceedings of the American Statistical Association meetings, 2008.

Meekins, Brian. "Possibilities to Improve Calling Efficiency in the American Time Use Survey." Bureau of Labor Statistics Working Paper, March, 2013.

Meekins, Brian and Stephanie Denton. "Cell Phones and Nonsampling Error in the American Time Use Survey," published in the proceedings of the American Statistical Association meetings, 2012.

McGuckin, Nancy. U.S. Department of Transportation, National Personal Transportation Survey. Personal Communication. 3 Aug. 1996.

National Academy of Sciences (NAS) and Committee on National Statistics, National Research Council. Time-Use Measurement and Research. Washington, D.C.: National Academy Press, 2000.

Nordhaus, William. "Time-Use Surveys: Where Should the BLS Go From Here?" Summary of the Conference on Time Use, Non-Market Work, and Family Well-Being, Hatch, Lynn (Ed.), Washington, DC: BLS and the MacArthur Network on the Family and the Economy, 1997.

O'Neill, Grace and John Dixon. "Nonresponse bias in the American Time Use Survey," published in the proceedings of the American Statistical Association meetings, 2005.

Robinson, John P. and Geoffrey Godbey. Time for Life: The Surprising Ways Americans Use Their Time. University Park, PA: The Pennsylvania State University Press, 1997.

Robison, E. (1999) "Sampling and Reporting in Time-Use Surveys," published in the proceedings of the American Statistical Association meetings.

Schwartz, Lisa K. (February, 2001). "Minding the Children: Understanding how recall and conceptual interpretations influence responses to a time-use summary question."

Schwartz, Lisa K., Jayme Gortman, and Siri Lynn. (July, 2001). "What's work? Respondents' interpretations of work-related summary questions."

Shelley, Kristina (2005). "Developing the American Time Use Survey activity classification system," *Monthly Labor Review*, June 2005, pp. 3-15.

Stewart, Jay (2002). "Assessing the Bias Associated with Alternative Contact Strategies in Telephone Time-Use Surveys." *Survey Methodology*. Volume 28, No. 2/December 2002, pp. 4-15.

American Time Use Survey (ATUS)

1220-0175

October 2016

Stinson, Linda. (2000). Final report of cognitive testing for the American Time Use Survey. Bureau of Labor Statistics internal report, August 2000.

Westat. "Research Services for Usability Testing and Lexicon Evaluation: The American Time Use Survey." October, 2001.