Contact Strategies and Activity Bias in Telephone Time-Use Surveys

Jay Stewart, Bureau of Labor Statistics, 2 Massachusetts Avenue, NE Room 4945

I. Introduction

The American Time-Use Survey (ATUS), which is currently under development at the U.S. Bureau of Labor Statistics (BLS), is the first time-use survey sponsored by the U.S. government since the studies conducted by the U.S. Department of Agriculture in the 1920s and 1930s. Beginning in January of 2003, the ATUS will collect time-diary data from approximately 1,000 individuals per month. To collect data from such a large sample while keeping costs under control, the BLS has decided to collect data via telephone. Ideally, all respondents will be available on their original calling day and will readily provide information to the interviewer. But this is unrealistic, so it is essential that ATUS and other telephone time-use surveys have a strategy for making subsequent attempts to contact respondents.

What is a contact strategy? A contact strategy is comprised of a contact schedule and a field period. The contact schedule specifies which days of the week that contact attempts will be made, and the field period specifies the maximum number of weeks attempts that will be made.

Designated-day and Convenient-day Schedules

Contact schedules fall into two main categories: designated day and convenient day. Both types of schedule randomly assign each respondent to a specific calling day. If the respondent is contacted the interviewer attempts to collect information about the reference day, which is the day before the calling day. However, the schedules differ in how interviewers make subsequent contact attempts.

Under a designated-day schedule, there are two approaches to making subsequent contact attempts. The interviewer could call the respondent on a later date, and ask the respondent to report activities for the original reference day. This approach maintains the original reference day, but extends the recall period. Harvey (1993) recommends allowing a recall period of no more than two days. The second approach is to postpone the interview and assign the respondent to a new reference day. Kalton (1985) recommends postponing the interview by exactly one week, so that the new reference day is the same day of the week as the original reference day.

These approaches are not mutually exclusive. For example, Statistics Canada's designated-day schedule allows interviewers to call respondents up to two days after the reference day, and to postpone the interview by one week if the respondent cannot be reached after the second day of attempts (see Statistics Canada 1999). The interview can be postponed up to three times.

The convenient-day schedule does not maintain the designated reference day. If no contact is made, then the interviewer calls on the next day and each subsequent day until the respondent is contacted. Once contact is made, the interviewer attempts to collect the interview or, if the respondent is unwilling to complete the interview at that time, reschedule it to a day that is convenient for the respondent. The reference day is always the day prior to the interview. It is worth noting that because respondents are not likely to schedule interviews on busy days, allowing them to choose their interview day is really no different than the interviewer proposing consecutive days (or calling on consecutive days) until the potential respondent accepts. Hence, one may think of the convenient-day schedule as being functionally identical to an every-day contact schedule.

Most methodological papers argue in favor of using a designated-day schedule (Kinsley and O'Donnell 1983, Kalton 1985, Lyberg 1989, and Harvey 1993 and 1999). Lyberg (1989) argues that the convenient-day schedule may introduce bias because "the respondent may choose a day when he/she is not busy, a day he/she is not engaged in socially acceptable behavior, a day he/she thinks is representative, etc." And Kinsley and O'Donnell (1983) argue that the convenient-day schedule could exaggerate the number of events taking place outside the home, because the respondent is more likely to be interviewed on a day that immediately follows a day that he or she was out of the house.

Two of these studies directly compare designated-day and convenient-day schedules (Kinsley and O'Donnell 1983 and Lyberg 1989). In the Kinsley and O'Donnell study, the experimental design divided the sample into two groups. They found that the two schedules produced similar response rates, and that the demographic composition of both samples were similar. However, they also found that the estimated time spent away from home was much higher under the convenient-day schedule than under the designated-day schedule. In the Lyberg study, two diaries were collected from each respondent. One was collected using a designated-day schedule and the other was collected using a convenient-day schedule. However, the convenient-day diaries were conducted by an interviewer, while the designated-day diaries were self-administered several days after the convenient-day interview. So it is impossible to determine whether any differences were due to differences in contact schedules or whether they were due to mode effects.

Two studies (Lyberg 1989 and Laaksonen and Pääkkönen 1992) investigate postponement. Both find that postponement increases response rates, but does not introduce bias into the estimates.

One advantage of the convenient-day schedule is that it is possible to make many contact attempts in a short period of time. In contrast, the designated-day schedule--as proposed--permits only one contact attempt per week. So it is natural to ask: Would it be reasonable to modify the designated-day schedule to allow some form of day-of-week substitution? For example, if the respondent cannot be reached on Tuesday to report about Monday, would it be acceptable to contact the respondent on, say, Thursday and ask the respondent to report about Wednesday? This modified schedule would allow for more contact attempts without having to extend the field period.

Because this type of substitution makes sense only if the substitute days are fairly similar to the original days, the first step was to determine which days, if any were similar to one another. In earlier work (Stewart 2000), I showed that Monday through Thursday are very similar to each other. Fridays are slightly different from the other weekdays. And Saturday and Sunday are very different from the weekdays and from each other. The five weekdays are similar enough to each other to permit day-of-week substitution. However, because Friday diaries are collected on Saturdays, which is the easiest day of the week to contact respondents, Fridays would likely be oversampled if they were included in the substitution schedule. Therefore, BLS is considering day-of-week substitution only for Monday through Thursday diary days (Tuesday through Friday contact days).

When selecting a contact strategy, we need to be concerned with two types of bias: activity bias and noncontact bias. Activity bias occurs when the probability of contacting and interviewing a potential respondent on a particular day is correlated with the respondent's activities on that diary day. Note that here and throughout the paper, the term contact probability refers to the probability of a productive contact (one that results in an interview). Noncontact bias occurs when contact probabilities are correlated with activities across individuals. Because noncontact bias has been examined extensively, I focus on activity bias.[1] A simple numerical example will illustrate this bias.

Example of Activity Bias: Suppose that all potential respondents are identical, and that their days fall into two categories: hard-to-contact (HTC) days and easy-to-contact (ETC) days. Further suppose that half of each potential respondent's days are HTC, half are ETC, and that HTC and ETC days are randomly distributed over the calendar. Finally, suppose that interviewers never contact respondents on HTC days (i.e., that $P_H = 0$, where P_H is the probability of contacting the respondent on an HTC day), and that they always contact respondents on ETC days (i.e., that $P_E = 1$). For simplicity, I assume that the activities of a given day can be summarized by an "activity index," I_J, where $I_J = 1 - P_J$ ($J = H,E$). The activity index represents time spent in activities that are negatively correlated with the contact probability--for example, activities done away from home). The average true activity index for this sample of potential respondents is 0.5 ($= 0.5 \times 1 + 0.5 \times 0$).

If a convenient-day contact schedule is used, then HTC days are oversampled. To see why this occurs, it is instructive to work through the two possible contact sequences. If the initial contact attempt occurs on an ETC day, then the respondent is contacted and asked about the previous day (the diary day). Because there are equal numbers of HTC and ETC days and these days are randomly distributed over the "calendar," half of these diary days will be HTC and the other half will be ETC. Therefore, the average activity index for the diary days of these respondents equals 0.5, which is the same as the population average. If, on the other hand, the initial contact day is an HTC day, then no interview takes place and the respondent is called on the following day. Contact attempts continue every day until contact the respondent is reached (on an ETC day). The average activity index for the diary days of these respondents equals 1, because the respondent is always interviewed on an ETC day that immediately follows an HTC day. Since half of the initial contact attempts are made on HTC days and the other half are made on ETC days, the average activity index for the final sample is equal to 0.75 ($= 0.5 \times 0.5 + 0.5 \times 1$).

Activity bias is not limited to time-use surveys. For example, in addition to their main focus on collecting event history information on employment, the National Longitudinal Surveys also include a few questions about employment and hours during the week prior to the interview. Because these interviews tend to be scheduled at the convenience of the respondent, the activities of the reference week will be correlated with the probability of interviewing the respondent about that reference week. This correlation introduces bias into hours-worked estimates, although the direction of the bias is indeterminate. Hours are overestimated for respondents who were unable to schedule an interview because of a heavy work schedule, and are underestimated for respondents who were away on vacation. Activity bias is also an issue for travel surveys. Time spent away from home will tend to be overestimated if respondents are asked about, say, the

[1] For a more complete discussion, see Stewart (2001).

four weeks prior to the interview. Asking respondents about a fixed reference period can eliminate this bias.

The rest of the paper examines the activity bias associated with four different contact schedules under alternative assumptions about the correlation of activities across days of the week. In Section II, I introduce the four schedules, and use simulations to assess the bias associated with each schedule. Section III concludes and makes recommendations.

II. Contact Schedules, Correlated Activities, and Activity Bias

In this section, I compare the activity bias associated with the convenient-day schedule and each of the three variants of the designated-day schedules. These schedules are defined as follows:

1. Convenient-day (CD): Attempt to contact potential respondents every day following the initial contact attempt until the potential respondent is contacted or until the field period ends.
2. Designated-day (DD): Attempt to contact respondents only once (no subsequent attempts).
3. Designated-day with postponement (DDP): Attempt to contact potential respondents on the same day of the week as the initial attempt until the potential respondent is contacted or until the field period ends (Kalton 1985).
4. Designated-day with postponement and substitution (DDPS): Attempt to contact potential respondents every other day following the initial contact attempt until the potential respondent is contacted or until the field period ends.

The DDPS schedule assumes alternating Tuesday/Thursday and Wednesday/Friday contact days. Whether the first week is Tuesday/Thursday or Wednesday/Friday depends on the start day, which is randomly assigned.

As we saw in the example above, a convenient-day schedule introduces activity bias into time-use estimates when activities are uncorrelated across days. Designated-day strategies should eliminate this bias. However, when activities are correlated across days,[2] it is considerably more complicated to determine whether sample estimates are biased, and to determine the direction and extent of that bias. Computer simulations are an ideal way to assess the bias associated with these contact strategies under alternative assumptions about the correlation of activities across days.

Simulations

My simulation strategy was very straightforward. First, I created four weeks worth of "data" for each of 10,000 potential respondents.[3] Because the simulations are designed to compare the four contact schedules above, I assumed that the "week" is five days long and that the eligible diary days are Monday through Thursday. The next step was to simulate attempts to contact these respondents using the four schedules described above.[4] Finally, I compared the estimates generated using each schedule to the true values.

To simplify the simulations I abstracted from specific activities, as in the example above, and characterized each day using an activity index, I_J, (J = H,E) that ranges from 0 to 1. The probability of contacting and interviewing the respondent on any given day is equal to $P_J = 1 - I_J$ (J = H,E). I assume that $P_H < P_E$, which means that respondents are less likely to be contacted on HTC days than on ETC days. To simulate the variation in activities across days, the contact probability on a given day is given by:

$$P_J = \overline{P}_J + \varepsilon ,$$

where $\overline{P}_J$ is the average contact probability on an HTC (J = H) or an ETC (J = E) day, and $\varepsilon \sim U(-\hat{\varepsilon},\hat{\varepsilon})$. To insure that contact probabilities lie in the [0,1] interval, I assume that $\hat{\varepsilon} < \min(\overline{P}_H, 1-\overline{P}_E)$.

There are many assumptions one can make regarding the correlation of activities across days. The simplest case is where activities are not correlated across days. However, it is possible that potential respondents are systematically harder to contact on some days than others. For example, some respondents may be hard to contact at the beginning of the week or on certain days of the week. Unfortunately, there is no way to be certain how activities are correlated across days. So for this reason, I ran the simulations under following eight assumptions about the pattern of HTC and ETC days in each of the four weeks:

1. Actual values of the activity index are distributed as U(0,1), so that the average value is 0.5.
2. The first two days of every week are HTC and the last three days are ETC (HHEEE).

---

[2] Activities are correlated, for example, if the respondent is busy on Monday, Wednesday, and Friday of each week (negative serial correlation), or if the respondent is busy for the first half of each week (positive serial correlation).

[3] Because I am focusing on contact strategies, I ignore the sampling procedures and assume that the sample of potential respondents is representative of the population.

[4] A copy of the program used to generate the sample and simulate contact attempts is available from the author on request.

3. The first three days of every week are HTC and the last two days are ETC (HHHEE).
4. The first four days of every week are HTC and the last day is ETC (HHHHE).
5. The first day of every week is ETC and the last four are HTC (EHHHH).
6. The first two days of every week are ETC and the last three are HTC (EEHHH).
7. The first three days of every week are ETC and the last three are HTC (EEEHH).
8. For half the sample Monday, Wednesday, and Friday are HTC and Tuesday and Thursday are ETC (HEHEH). For the other half of the sample the reverse is true (EHEHE).

In pattern 1, activities are uncorrelated across days. Patterns 2-7 correspond to positive serial correlation, and pattern 8 corresponds to negative serial correlation.

Table 1 shows the results from a representative subset of the 153 simulations I ran. The first four columns show the average contact probability on HTC and ETC days, the value of $\hat{\varepsilon}$, and the true average activity index. The remaining columns show the estimated average activity index for each of the four contact schedules. Shaded entries indicate that the estimated average is statistically different from the true average (at the 5 percent level).[5]

Pattern 1 - No Correlation

This pattern is essentially the same as in the numerical example above. The main result is that all of the contact strategies generate unbiased estimates for the average activity index, except the CD strategy. As expected, the CD strategy overestimates the average activity index. More importantly, when using the CD strategy, the estimated average activity index--and hence the bias when activities are uncorrelated across days--*is positively correlated with the variance of ε.* One can see the intuition behind this result by noting that a large negative realization of ε on a particular day makes it less likely that the respondent will be contacted on that day, and hence, more likely that that day will become the diary day. None of the other contact strategies are sensitive to the variance of ε.

Patterns 2-7 - Positive Serial Correlation

The results are mixed when activities exhibit positive serial correlation across days of the week. The CD and DDP schedules generate estimates that are closest to the true means under patterns 2 (HHEEE) and 3 (HHHEE). Under patterns 4 (HHHHE) and 5 (EHHHH), all of the designated-day schedules perform well. Only the DDP schedule exhibits any statistically significant bias (for the EHHHH pattern), and the bias is rather small. The DDP schedule is closest to the true value under schedules 6 (EEHHH), and 7 (EEEHH). However, the other schedules perform reasonably well when the difference between P_H and P_E is small. As above, the estimated average activity index increases with the variance of ε under the CD schedule, but not under any of the other schedules.

Pattern 8 - Negative Serial Correlation

All of the contact schedules generate biased estimates, because ETC days are undersampled. The DDP schedule generates the smallest bias. The DDPS schedule generates a large activity bias, because contact attempts are made on two HTC days and then on two ETC days (or the reverse). This pattern results in contacting respondents on a relatively large fraction of ETC days, so that diary days are disproportionately HTC days. Not surprisingly, if the DDPS schedule is modified so the respondent is contacted on the same two days each week, there is virtually no bias.

It is clear from these simulations that the activity bias associated with each contact strategy depends on the correlation of activities across days, the contact probabilities on HTC and ETC days, and the variance of those probabilities.

Correlation of Activities Across Days

To gain some insight about the correlation of activities across day of the week, I used data from the May 1997 Work Schedule Supplement to the Current Population Survey (CPS) to tabulate frequency of various work schedules. Table 2 shows the patterns of work days and nonwork days from the May 1997 CPS. Note that because I am interested in the prevalence of each type of schedule for the entire population, Table 2 includes both workers and nonworkers. Nearly 88 percent of all individuals fall into two patterns: 48 percent work all five weekdays, and 39 percent work no weekdays. Another 4 percent work four weekdays and have either Friday or Monday off. The remaining individuals exhibit no discernible pattern. Hence, contact probabilities of the vast majority of individuals are roughly the same each day except for random noise, which corresponds to the no-correlation pattern above.

IV. Conclusions and Recommendations

Contacting respondents in telephone time-use surveys presents some unique challenges. Unlike most

[5] I chose N so that so that most deviations from the true averages would not be due to random noise. Standard errors for these simulations range between 0.0003 and 0.006.

Table 1: Activity bias associated with each contact strategy under alternative assumptions about the correlation of activities across days

	Average Contact Probability				Estimated Activity Index			
Activity Pattern	Hard-to-contact days	Easy-to-contact days	ε-hat	True Average Activity Index	CD	DD	DDP	DDPS
No Serial Correlation								
	0.500		0.100	0.500	0.503	0.499	0.500	0.500
	0.500		0.300	0.500	0.526	0.498	0.500	0.501
	0.500		0.500	0.500	0.575	0.495	0.502	0.503
Positive Serial Correlation								
HHEEE	0.750	0.250	0.050	0.500	0.503	0.447	0.476	0.431
	0.750	0.250	0.250	0.500	0.526	0.445	0.476	0.431
	0.600	0.400	0.050	0.500	0.499	0.489	0.496	0.486
	0.600	0.400	0.200	0.500	0.512	0.487	0.496	0.488
HHHEE	0.750	0.250	0.050	0.625	0.608	0.565	0.600	0.545
	0.750	0.250	0.250	0.625	0.630	0.560	0.599	0.545
	0.600	0.400	0.050	0.550	0.548	0.540	0.547	0.536
	0.600	0.400	0.200	0.550	0.560	0.537	0.547	0.538
HHHHE	0.750	0.250	0.050	0.750	0.7504	0.7493	0.7504	0.7501
	0.750	0.250	0.250	0.750	0.7670	0.7464	0.7510	0.7514
	0.600	0.400	0.050	0.600	0.6006	0.5997	0.6003	0.6003
	0.600	0.400	0.200	0.600	0.6113	0.5979	0.6009	0.6011
EHHHH	0.750	0.250	0.050	0.625	0.6355	0.6315	0.6338	0.6296
	0.750	0.250	0.250	0.625	0.6511	0.6231	0.6320	0.6294
	0.600	0.400	0.050	0.550	0.5558	0.5519	0.5529	0.5519
	0.600	0.400	0.200	0.550	0.5657	0.5497	0.5533	0.5521
EEHHH	0.750	0.250	0.050	0.500	0.409	0.415	0.479	0.392
	0.750	0.250	0.250	0.500	0.420	0.410	0.477	0.395
	0.600	0.400	0.050	0.500	0.490	0.489	0.498	0.487
	0.600	0.400	0.200	0.500	0.498	0.488	0.499	0.487
EEEHH	0.750	0.250	0.050	0.375	0.313	0.309	0.355	0.299
	0.750	0.250	0.250	0.375	0.332	0.309	0.354	0.301
	0.600	0.400	0.050	0.450	0.441	0.440	0.448	0.439
	0.600	0.400	0.200	0.450	0.450	0.439	0.448	0.439
Negative Serial Correlation								
HEHEH/ EHEHE	0.750	0.250	0.050	0.500	0.658	0.632	0.548	0.643
	0.750	0.250	0.250	0.500	0.673	0.632	0.549	0.647
	0.600	0.400	0.050	0.500	0.528	0.523	0.506	0.526
	0.600	0.400	0.200	0.500	0.539	0.522	0.506	0.525

Note: Shading indicates that the estimated average activity index is statistically different from the true value at the 5% level.

other surveys, time-use surveys cannot accept proxy responses, so it is more difficult insure that the final sample is representative. And because telephone time-use surveys ask respondents to report on their activities during the previous day, it is essential to insure that there is no correlation between the probability of contacting the potential respondent and his or her activities on the diary day.

Running a set of simple simulations, I have shown that the choice of contact strategy matters, and that the bias associated with any given strategy depends on the correlation of activities across days. None of the strategies produced unbiased results under all possible assumptions about the correlation of activities across days, so the optimal strategy depends on the incidence of each type of correlation in the population. Direct data on these correlations do not exist. But data from the May 1997 Work Schedule Supplement to the CPS suggest that the contact probabilities of the vast majority of individuals are likely to be uncorrelated across days.

The recommendations are straightforward. First, time-use surveys should use a designated-day contact strategy. And second, because there does not seem to any difference between the DDP and DDPS strategies, survey managers can use other criteria, such as cost, to select a contact strategy.

References

Harvey, Andrew (1993) "Guidelines for Time Diary Data Collection," Social Indicators Research 30, pp. 197-228.

Harvey, Andrew (1999) "Guidelines for Time Use Data Collection and Analysis," in Wendy E. Pentland, Andrew S. Harvey, Powell Lawton, and Mary Ann McColl (Eds.) *Time Use Research in the Social Sciences*, pp. 19-45, edited by. Kluwer Academic/Plenum Publishers, New York.

Kalton, Graham (1985) "Sample Design Issues in Time Diary Studies" In F. T. Juster & F. T. Stafford (Eds.), Time, Goods, and Well-Being, pp. 333-351. Ann Arbor: University of Michigan, Institute of Social Research.

Kinsley, Brian and Terry O'Donnell (1983) "Marking Time: Methodology Report of the Canadian Time Use Pilot Study--1981." In *Explorations in Time Use* (vol. 1), Ottawa: Department of Communications, Employment and Immigration.

Laaksonen, Seppo and Hannu Pääkkönen (1992) "Some Methodological Aspects on the Use of Time Budget Data." In *Housework Time in Bulgaria and Finland*, edited by L. Kirjavainen, B. Anachkova, S. Laaksonen, I. Niemi, H. Pääkkönen, Z. Staikov, pp.86-104.

Lyberg, Ingrid (1989) "Sampling, Nonresponse, and Measurement Issues in the 1984-85 Swedish Time Budget Survey." In *Proceedings of the Fifth Annual Research Conference*: Department of Commerce, Bureau of the Census, pp. 210-238.

Statistics Canada (1999) *Overview of the Time Use of Canadians in 1998, General Social Survey*, Catalogue no. 12F0080XIE; Ottawa, Canada.

Stewart, Jay (2000) "Alternative Indexes for Comparing Activity Profiles," paper presented at the 2000 International Association for Time-Use Research Conference, Belo Horizonte, Brazil.

Stewart, Jay (2001) "Assessing the Bias Associated with Alternative Contact Strategies in Telephone Time-Use Surveys." Unpublished paper Bureau of Labor Statistics.

Triplett, Timothy (1995) "Data Collection Methods for Estimating Exposure to Pollutants Through Human Activity Pattern Data: A National Micro-behavioral Approach," mimeo, Survey Research Center, University of Maryland.

Table 2: Distribution of work schedules

Activity Pattern					Percent	Cumulative Percent
M	Tu	W	Th	F		
-	-	-	-	-	39.40	39.40
W	W	W	W	W	48.11	87.51
W	W	W	W	-	2.63	90.14
-	W	W	W	W	1.63	91.77
W	W	W	-	-	0.81	92.58
W	W	-	-	-	0.26	92.84
-	-	-	W	W	0.37	93.21
-	-	W	W	W	0.68	93.89
W	-	W	-	W	0.49	94.38
-	W	-	W	-	0.25	94.63
-	-	-	-	W	0.51	95.14
W	-	-	-	-	0.25	95.39
W	W	-	W	W	0.73	96.12
W	-	-		W	0.36	96.48
W	-	-	W	W	0.70	97.18
Other patterns					2.82	100.00
Total					100.00	

Note: A "W" indicates a workday, and a "-" indicates a nonwork day. Author's tabulations from the May 1997 Work Schedule Supplement to the CPS. Observations were weighted using supplement weights. N = 89746