Measuring intrahousehold allocation of time: response to Anne E. Winkler

Measuring the use of time by more than one individual in a household, though important, cannot be accomplished within the data quality requirements and budgetary constraints of the new BLS American Time Use Survey; the topic, however, is on the Bureau’s agenda for 2002.

The current design for the BLS American Time Use Survey does not include questions intended to measure the use of time by household members other than the individual who is selected as the “designated person.” In an article in this issue of the Review, Anne E. Winkler recommends that the United States adopt an approach similar to one used by Statistics Canada to obtain some information about time use by more than one individual in a household. The designers of the American Time Use Survey, as well as many others, agree that measuring all adult household members’ use of time would produce valuable information on how household members jointly— and not just individuals singly—allocate their time. However, a conscious decision was made not to attempt to collect this kind of information in the first year of full production of the survey, chiefly because it was necessary to devote developmental resources to other issues.

Two primary considerations guided the design of the survey: the needs of prospective data users and the feasibility of conducting a time-use survey by telephone. A survey of the time-use literature and a review of papers presented at two conferences on time use were conducted to determine user interests. The measurement of market work and that of nonmarket work emerged as frequent areas of interest, with child care identified as a nonmarket activity of principal concern to many data users. The Bureau realized that, given the immense amount of work needed to field a new, ongoing time-use survey and the limited resources available, not all prospective data users’ interests could be met. On the basis of the literature review and the identified needs of data users, it was decided to focus development efforts on disambiguating work activities and collecting comprehensive data on child care as both a primary and a secondary activity.

Because the designers of the American Time Use Survey agree with Winkler’s suggestion that measuring the intrahousehold allocation of time is very important, and because the survey will not be collecting the relevant data in the first year, a considered response to Winkler’s suggestion is merited. Thus, the objective of this article is twofold: to offer an update on the design of the survey and to discuss some issues related to the measurement of time use by more than one individual in a household.

Briefly, the survey will interview a randomly selected individual from a subset of households that complete month-in-sample 8 (MIS8) of the Current Population Survey (CPS). Respondents to the American Time Use Survey will be interviewed one time only, about the previous day’s activities. The Bureau of the Census will conduct the survey as a computer-assisted telephone interview. The main component of the survey will be its core time diary, which covers a 24-hour period from 4:00 A.M. the previous morning until 4:00 A.M. the day of the interview. Upon completion of the core time diary, summary questions will be asked which collect additional information that may not be reported in the diary. Based on the results of a feasibility test conducted in 1997 and an operations field test conducted in April–June.
2001, it was decided that the survey would consist of a telephone interview with no field followup. The fielding period will be 8 weeks, with a new sample introduced weekly. Respondents will be assigned one day per week on which they will be eligible for an interview.

Summary of Winkler’s article

Neither the initial plans nor the current plans for the American Time Use Survey include conducting interviews with more than one member of the same household or allowing proxy reports of time use. Several other countries that conduct time-use surveys have collected full diary information from several members of a household. Most of these surveys have used self-administered paper diaries to collect information on time use. In her article, Winkler suggests that the United States adopt a middle-ground approach to obtaining information about time use by more than one member of a household. Specifically, she recommends following the Canadian model: in 1998, Statistics Canada conducted time-use interviews by telephone with one respondent per household and then collected some information about spouses’ or unmarried opposite-sex partners’ time use through a series of “stylized” questions. For example, Statistics Canada asked respondents, “Last week, did [your spouse or partner] do any unpaid work to maintain or improve your house, yard, or automobile?” Positive responses were followed up with a question that asked respondents to estimate how many hours their spouse or partner engaged in those activities. The use of a series of questions of this type allows for some proxy reporting of time use without requiring the level of detail that is obtained in the respondent’s own 24-hour recall diary.

Winkler’s comments about the American Time Use Survey appear to be based chiefly on her review of a National Research Council report on a time-use measurement and research workshop that was held in May 1999. Some of her recommendations were implemented in the most recent plans for the American Time Use Survey, while others were not. Winkler’s central point—that measuring the use of time by more than one member of a household is important—is echoed by others as well. First, collecting time-use information from more than one member of a household could shed more light on the labor force participation of all members of the household. Specifically, she recommends following the Canadian model: in 1998, Statistics Canada conducted time-use interviews by telephone with one respondent per household and then collected some information about spouses’ or unmarried opposite-sex partners’ time use through a series of “stylized” questions. For example, Statistics Canada asked respondents, “Last week, did [your spouse or partner] do any unpaid work to maintain or improve your house, yard, or automobile?” Positive responses were followed up with a question that asked respondents to estimate how many hours their spouse or partner engaged in those activities. The use of a series of questions of this type allows for some proxy reporting of time use without requiring the level of detail that is obtained in the respondent’s own 24-hour recall diary.

Winkler’s comments about the American Time Use Survey appear to be based chiefly on her review of a National Research Council report on a time-use measurement and research workshop that was held in May 1999. Some of her recommendations were implemented in the most recent plans for the American Time Use Survey, while others were not. Winkler’s central point—that measuring the use of time by more than one member of a household is important—is echoed by others as well. First, collecting time-use information from more than one member of a household could shed more light on the labor force participation of all members of the household. Second, such data could afford more information on the division of time and tasks within households—in particular, information about who, within the household, provides care for which other members of the household.

Collecting data on more than one member of a household

Measurement issues. Many participants in the National Research Council workshop on time-use measurement and research agreed that data should be collected from more than one member of a household. They also noted that several problems could arise with this approach. Response rates could suffer because interviewing a number of individuals would mean that the survey would take longer. Operational problems could arise due to difficulty in gaining cooperation from all of the members of the household who are to be interviewed on the same day, to provide information about the same day. Note that getting information from each individual about the activities they engaged in on the same day would be necessary to fully understand the joint allocation of time within a household.

The quality of the data could suffer due to an extended recall period for some respondents. It is generally assumed that the quality of data in time-use surveys suffers the longer the retrospective recall period. Therefore, most surveys ask respondents only about the previous day’s activities. In the event that two or more members of a household are unavailable on the same day, survey designers are faced with two possible courses of action: interviewers could conduct the second person’s interview at a later date, but still seek information about the original reference day, or the second respondent could be interviewed at a later date, with the interviewer seeking information about a different reference day. Unfortunately, neither of these strategies is ideal: the first introduces recall error, while the second reduces the reliability of data with respect to analyzing the intrahousehold allocation of time.

In addition to this pair of measurement issues, a number of other important methodological considerations must be weighed regarding the collection of time-use information from more than one member of a household. The next subsection discusses other countries’ experiences with administering multiple diaries in a household.

Paper diaries versus retrospective interviews by telephone. Each survey that has collected time-use data from more than one member of a household has relied on paper diaries, supplemented by household or individual-level self-administered questionnaires or personal visits. The choice of paper diaries, which are essentially prospective (for example, with instructions such as “Tomorrow, write down all your activities”), versus retrospective interviews by telephone (for example, with questions such as “What did you do yesterday?”) has important implications for response rates and survey costs. When paper diaries are combined with personal visits to drop off the diaries and explain the process of keeping a diary, they afford interviewers an opportunity to assess the household’s likely compliance with the instructions before assigning all household members a designated day on which to keep their diaries. Eurostat’s pilot harmonized time-use survey procedures included postponement plans that
were implemented when one or more diarists were unable to keep the diary on the originally assigned day. This type of prescreening may not be a viable option with a telephone survey. In a recent American Time Use Survey field test, researchers tested the effectiveness of proactive appointment setting, in which interviewers called respondents in advance to schedule a time to complete the survey. The procedure failed to increase response rates, but did increase overall costs.\(^7\)

Statistics Canada tested the feasibility of conducting time-use interviews via telephone with more than one member of a household. In the test, a single household respondent was interviewed about two different days of the week, and the married partners of those respondents were interviewed about the same days of the week. As part of the test, diaries from spouses of the respondents were collected only from households that were asked to complete one weekend and one weekday interview, for a total of four time-use interviews per household. Statistics Canada found that the collection of time-use information from two household respondents increased the length of the interview by about 40 minutes. Approximately 16 of the 40 minutes were taken up by extra explanations and extra calls to reach the spouse, and about 24 minutes were used to conduct the time-diary interview with the spouse. In addition, the interviewers found it difficult to collect all of the spouses’ diaries: only 46 percent of those households completed all four interviews, compared with 88 percent from single-respondent households (that had multiple interviews). Statistics Canada estimated that if the requirement for completeness had been only two interviews—one with the respondent and one with the spouse, about the same day—the completion rate for both spouses would have increased to 63 percent. Given the low response rate for spouses and the high cost in time, Statistics Canada concluded that it was not advisable to collect this information.\(^8\)

As Winkler points out, in the United States the National Survey of Families and Households uses both personal interviews and self-administered questionnaires to collect information from more than one member of a household. However, the interviews and questionnaires are very long. In developing the survey, designers attempted to maintain an average interview length of 90 minutes. Of the interviews conducted between March 1987 and May 1988, the majority took between 70 and 110 minutes to complete.\(^9\) In contrast, the BLS survey is expected to be completed in approximately 30 minutes.

Two U.S. time-use surveys, the 1975–76 and 1981–82 surveys administered by the University of Michigan, collected multiple 24-hour-recall diaries from respondents and their spouses in separate interviews conducted at 3-month intervals. Both spouses were required to report about the same day. This design specification resulted in some cases that required respondents to recall activities that occurred more than 24 hours earlier. In general, the extended recall period was required when the interview with the spouse could not be obtained at the same time as that with the respondent and thus was conducted some days later. One researcher who examined the effects of the length of the recall period on the quality of time-use data, as measured by the number of primary activities reported, found that reports (about Monday through Thursday) included 10 percent to 20 percent fewer activities when the recall period extended beyond 24 hours.\(^10\)

Exhibit 1 summarizes data collection methods and response rates of several international time-use surveys. With the exception of Canada, the countries represented were chosen because they collect diaries from more than one member of a household. Canada is included because Statistics Canada’s

### Exhibit 1: Summary of time-use surveys

<table>
<thead>
<tr>
<th>Household survey and sponsoring agency</th>
<th>Sample</th>
<th>Methodology</th>
<th>Response rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSS 1998, Cycle 12, Time Use Survey, Statistics Canada</td>
<td>Sample of households obtained by random-digit dialing. One eligible respondent 15 years or older per household</td>
<td>Computer-assisted telephone interview of 24-hour-recall diary</td>
<td>78 percent at individual level</td>
</tr>
<tr>
<td>1997 Australia Time Use Survey, Australian Bureau of Statistics</td>
<td>All eligible household members 15 years or older in 4,100 households</td>
<td>Self-completed 24-hour diary</td>
<td>72 percent for all members of the household; 84 percent for one member of the household</td>
</tr>
<tr>
<td>1998/1999 Time Use Survey, Statistics New Zealand, Ministry of Women’s Affairs</td>
<td>National sample of 7,200 households. Two eligible respondents per household</td>
<td>Self-completed 48-hour paper diary, individual and household questionnaires, personal visits</td>
<td>72 percent for both members of the household</td>
</tr>
</tbody>
</table>
methodology is most similar to that planned for the American Time Use Survey.

The data from the 1997 Australian time-use survey and from Statistics Canada’s pretest indicate that household-level response rates are lower than individual-level rates, which suggests that interviewing more than one adult per household could have a negative impact on response rates in the United States.

Data from a 1985 time-use survey conducted jointly by the University of Maryland and the U.S. Environmental Protection Agency indicate that response rates are highest when paper diaries are combined with personal visits, but response rates are only slightly lower for retrospective telephone interviews. However, personal visits are expensive, and the additional gains in response may not offset the increases in survey costs.11

Using stylized questions

Cognitive burden. Winkler notes that the use of stylized questions can be problematic in part because respondents “must go through an extensive ‘cognitive process’ in order to answer [them].”12 The cognitive demands associated with answering these kinds of questions may be so burdensome that respondents resort to “satisficing,” rather than maximizing the quality of their responses.13 Thus, they may expend less cognitive effort thinking about the meaning of a question, they may not search their memories as thoroughly, they may not be as careful in integrating information they recall, or they may respond imprecisely. One author notes that, when survey interviews become lengthy and burdensome, even respondents who may wish to provide high-quality data may become fatigued and resort to satisficing.14 To compound the problem, the level of effort respondents expend in coming up with an answer may further decline as their level of motivation declines.

In another study, F. T. Juster and F. P. Stafford suggest that the nature of the recall task inherent in recollecting one’s use of time underscores why diaries are the “only valid measurement of time use.”15 In providing time-use information, respondents are asked to recall activities that are not particularly memorable, that do not recur on a daily basis, and that are not amenable to the use of market measurements as a proxy. Given the nature of the task, Juster and Stafford conclude that it is not possible to get valid estimates of actual time used from simple survey questions which ask about the respondent’s typical use of time over some specified period.

In his review of time-use surveys, R. Andorka noted that, while it is possible to collect time-use data with stylized questions, time diaries provide valid and reliable data and “ought to be preferred, in spite of higher costs, to the method of stylized questions.”16 His conclusions are consistent with previous research carried out by J. P. Robinson, who found that 24-hour recall diaries provided better data than did stylized survey questions which asked about time spent in various activities.17 In Robinson’s studies, 24-hour recall diaries were compared with (1) reports obtained when respondents wrote down their activities whenever a randomly programmed beeper signaled, (2) detailed descriptions of activities that occurred during a randomly selected hour, and (3) responses to stylized survey questions. The time diaries were not significantly different from either the reports prompted by the random beeper or the descriptions of random-hour activities and were superior to the responses to stylized questions. Table 1 summarizes Robinson’s findings.

Proxy reports. For certain kinds of observed characteristics, proxy reports may be as reliable as self-reports. Generally speaking, though, proxy reports are recognized as a less reliable source of information. Their reliability is especially likely to be called into question when a respondent is asked to report about other household members’ activities. In a recent cognitive test of the CPS computer and Internet use supplement, L. Schwartz and S. Fricke examined the accuracy of proxy reports through a data verification process.18 Accuracy was measured by comparing respondents’ proxy reports with household members’ reports of their own computer and Internet use, using identical questions. Rates were based on proxy verifications obtained from 21 respondents who reported about their spouses’ computer and Internet use. Accuracy rates ranged from 20 percent to 100 percent, with the least accurate proxy reports centering on questions related to other household members’ online activities and their use of computers at work. In response to a debriefing question that asked respondents how they knew what other household members did on the computer, most respondents said that they would have to have been present or been told in order to know definitively what someone else in the household did.

Winkler also expressed concern about the reliability of proxy reports. She noted that, even though husbands’ and wives’ mean estimates of their time spent in various activities match up, there may be little consensus between matched husband-wife pairs. In fact, the correlation between husbands’ and wives’ reports in the National Survey of Families and Households was only 0.37 for reports about husbands’ activities and 0.46 for reports about wives’ activities.

Winkler notes that egocentric bias tends to widen the gap in hours attributed to various activities in self- and proxy reports.19 According to some psychologists, egocentric bias occurs in part because information about the self is more deeply encoded in memory and is more readily available for retrieval. This is consistent with findings in cognitive psychology which suggest that self-referencing (that is, applying to the self materials that are to be learned)
enhances retention. Taken together, the evidence suggests that the use of stylized questions to collect self- and proxy reports of time use is likely to result in overestimation of self-reports of socially desirable behaviors and underestimation of others’ activities. Without also interviewing other household members, the American Time Use Survey would be unable to provide BLS analysts with reliable enough information to verify the accuracy of proxy reports and, as a result, might be unable to measure, to any significant degree of accuracy, the true division of resources among household members.

Varying recall periods. The current design of the American Time Use Survey focuses the respondent on the previous day’s activities. The time diary and the work-related and child-care summary questions all refer the respondent to the previous day, while the summary question about absences from home asks respondents to recall trips taken in the previous month. Asking respondents to focus on different periods throughout the interview may be a cognitively difficult task. The implementation of stylized questions requires a careful consideration of the appropriate reference period in order to ensure high-quality data and analytical relevance.

Measuring intrahousehold time allocation with the current survey methods

Because the survey sample will be from the CPS, which gathers demographic and labor market information for the entire household, analysts will have available a rich set of controls for household members other than the respondent. Some of this information will have been collected in the previous month, but the survey instrument will update most labor market information from the respondent, as well as updating the spouse’s labor force status and usual hours of work. The background information from the CPS will allow the American Time Use Survey to be used in the analysis of the intra-household allocation of time. Analysts will be able to examine mean hours of time spent in a given activity by members of households of a given type, as well as mean differences between household members with a given relationship to each other.

For instance, analysts will be able to examine the hours spent in leisure activities by each member of a married couple. From this information, they will be able to estimate the average difference in time spent in that activity between husbands and wives. To take a numerical example, if one finds that husbands have, on average, 5 hours of leisure activities per day and wives have, on average, 3 hours, the average difference in time spent in leisure activities is obviously 2 hours per day. Such an analysis can then be extended by investigating labor market and demographic characteristics of the household. One might examine the hours spent in leisure and differences in leisure activities between husbands and wives when both work full time and compare the findings with hours spent in leisure when the husband works full time and the wife works part time. Other examples include estimating the average difference in time spent in an activity between husbands and wives when both work full time and compare the findings with hours spent in leisure when the husband works full time and the wife works part time. Other examples include estimating the average difference in time spent in an activity between husbands and wives with a given wage rate for the wife or a given difference in wage rates between the husband and wife. Other adults in the household can also be incorporated into the analysis. Assuming a sufficient sample size, one can examine, for example, the average difference in time spent in child care between parents with and without a grandparent in the household, as well as the average time spent in child care by within-household grandparents.

Because of the absence of data on more than one member of the same household, there are, of course, limits on what kinds of analyses are possible. For example, aside from

<table>
<thead>
<tr>
<th>Selected activities</th>
<th>Experiment 1 (women)</th>
<th>Experiment 2</th>
<th>Experiment 3 (women)</th>
<th>Experiment 3 (men)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beeper   Diary Random hour Diary</td>
<td>Sylized</td>
<td>Diary</td>
<td>Sylized</td>
</tr>
<tr>
<td>Work for pay</td>
<td>9.3   14.4</td>
<td>22.2</td>
<td>23.9</td>
<td>14.9</td>
</tr>
<tr>
<td>Housework</td>
<td>21.4</td>
<td>18.5</td>
<td>10.6</td>
<td>13.9</td>
</tr>
<tr>
<td>Child care</td>
<td>8.6</td>
<td>7.1</td>
<td>2.7</td>
<td>3.6</td>
</tr>
<tr>
<td>Shopping</td>
<td>4.3</td>
<td>6.6</td>
<td>7.5</td>
<td>6.8</td>
</tr>
<tr>
<td>Entertainment</td>
<td>3.7</td>
<td>5.7</td>
<td>7.4</td>
<td>9.1</td>
</tr>
<tr>
<td>Active leisure</td>
<td>5.8</td>
<td>4.0</td>
<td>3.4</td>
<td>2.8</td>
</tr>
<tr>
<td>Watching television</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Reading</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Note: Dash indicates data not collected.

examining averages, analyzing the distribution of time use across households is impossible. Thus, one can estimate the mean, but not the median, difference in leisure time between husbands and wives. The absence of data on the time use of the respondent’s spouse also makes it difficult to know when spouses are performing the same activity at the same time. For instance, to return to a previous example, if men give child care for 3 hours and women for 5 hours, how many hours are children in their parent’s care? Without simultaneous data on both men and women, it is impossible to tell. However, some indication of joint activities will be available in the survey, because respondents will be asked who was with them during each activity. Note that the stylized questions proposed by Winkler also do not allow us to detect joint activities.

AS NOTED BY PARTICIPANTS IN THE NATIONAL RESEARCH COUNCIL WORKSHOP ON TIME-USE MEASUREMENT, “the availability of data on multiple persons would greatly enhance the value of such data for understanding household behavior.”23 The Bureau of Labor Statistics recognizes the importance of measuring the intrahousehold allocation of time. However, the Bureau also recognizes that these measures may be difficult to capture with the current structure of the American Time Use Survey within reasonable budgetary and data quality constraints. On the basis of a review of the literature, considerable research would need to be done before questions of the type suggested by Winkler could be implemented. The Bureau is exploring methods for measuring the intrahousehold allocation of time, and the topic is on the BLS research agenda for 2002.

Notes


2 Designers were interested in the feasibility of conducting the survey by phone mainly for two reasons. First, drawing the sample from, and linking back to, the Current Population Survey (CPS) would be facilitated if the American Time Use Survey were a telephone survey. The CPS will provide a wealth of demographic information about respondents to the time-use survey. Second, as noted in F. T. Juster and F. P. Stafford, “The Allocation of Time: Empirical Findings, Behavioral Models, and Problems of Measurement,” Journal of Economic Literature, June 1991, pp. 471–522, telephone surveys cost a great deal less than personal-interview surveys without compromising the quality of time diary data.

3 The core time diary will be a listing of the previous day’s activities and associated contextual information. All citations of the diary in this article refer to the listing of activities collected during the telephone interview and do not refer to a paper-and-pencil diary completed by the respondent.

4 Stylized questions are survey questions that ask respondents to indicate how much time they spent in various activities or to estimate how often they engage in various activities over a predetermined reference period (See R. Andorka, “Time Budgets and Their Uses,” Annual Review of Sociology, vol. 13, 1987, pp. 149–64.)


6 J. Martin and R. Breten, “The Effect of Interviewer Characteristics on Survey Response Rates,” paper presented at the International Conference on Survey Nonresponse, Portland, or, October 1999, found that household surveys which interviewed only one responsible adult per household took an average of 36 minutes to complete. In comparison, household surveys that interviewed all household adults took an average of 85 minutes to complete.


9 Refusals to participate in the survey and breaking off the interview may be more common with telephone interviews than personal visits mainly because it is easier to put the phone down than it is to refuse someone calling in person. With this in mind, survey designers suggest that telephone interviews should not last longer than 20 minutes. However, shorter or longer interviews may be advised, depending on respondents’ level of interest in the survey topic. (See R. Thomas and S. Purdon, “Telephone methods for social surveys,” Social Research Update, 1994; on the Internet at www.soc.surrey.ac.uk/sru/sru8.html.

10 F. T. Juster, “Response Errors in the Measurement of Time Use,” Journal of the American Statistical Association, June 1986, pp. 390–402. Data showed that recall periods of longer than 24 hours resulted in reports that mentioned 5 percent fewer activities on Friday through Sunday. However, the difference was not statistically significant.

11 In the American Time Use Survey field test, a 74-percent response rate was obtained after 8 weeks of data collection for households with telephones. In comparison, after a total of 8 weeks, a 79-percent response rate was achieved with households that received a personal visit after 4 weeks of first trying to contact the respondent by phone. Despite this increase in response rate, a field component for the survey was cost prohibitive: it was estimated that the Bureau would incur an additional cost of approximately $102 per case if the survey included a field component.


15 Juster and Stafford, “The Allocation of Time.”


19 Egocentric bias is the tendency to be more sensitive to, and judge
differently, events involving oneself than events involving others.


22 Note that the analyst will have information on the respondent’s wages as of the previous CPS month if the respondent has not changed jobs and as of the current American Time Use Survey month if he or she has changed jobs. The analyst will have information on the spouse’s wages as of the previous CPS month.

23 Ver Ploeg et al., *Time Use Measurement and Research*, p. 49.

---

### Fax on demand

Users of data from the Bureau of Labor Statistics can request a fax of news releases, historical data, and technical information 24 hours a day, 7 days a week, from the Bureau’s fax-on-demand system.

Users can receive news releases of major economic indicators (see schedule on back cover) at 8:45 a.m. on the morning the data are released. The number to obtain data from the national office is:

**(202) 691–6325**

Use a touch-tone telephone and follow the voice instructions for entering document codes and your fax telephone number. The fax-on-demand catalog, containing a list of available documents and codes, can be obtained by entering code 1000. You may request up to four documents with each call. Faxes are sent immediately following the request. If your fax line is busy, the system attempts to send the requested material four times before disconnecting.