TECHNICAL WORKING PAPER SERIES

TIME USE AND POPULATION REPRESENTATION IN THE SLOAN STUDY OF ADOLESCENTS

Casey B. Mulligan
Barbara Schneider
Rustin Wolfe

Technical Working Paper 265
http://www.nber.org/papers/T0265

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2000

We appreciate the research assistance of Ananth Ramanarayanan and Jaeki Jeong. The views expressed in this paper are those of the authors and not necessarily those of the National Bureau of Economic Research.

Time Use and Population Representation in the Sloan Study of Adolescents
Casey B. Mulligan, Barbara Schneider, and Rustin Wolfe
NBER Technical Working Paper No. 265
November 2000
JEL No. C42, J22, J13

ABSTRACT

Do studies of time use interfere too much in the lives of the subjects? As a result are those who agree to participate a biased sample of the population? We examine the characteristics of the Experience Sampling Method (ESM) adolescent sample from the Alfred P. Sloan Study of Youth and Social Development in order to detect and quantify instances of sampling and nonresponse bias. According to available proxies for time use and standard demographic variables, the Sloan ESM sample is nearly representative in terms of teen employment rates, parental employment rates, a student's grade point average, and TV watching. Work hours are slightly undercounted in the study because of slightly higher nonresponse rates by teenagers working long hours. The sample is less representative in terms of the time of week and gender; nonresponse is relatively common on school nights and (to a lesser extent) on weekends, and among boys. We offer some suggestions regarding general implications of our findings for the measurement of time use.

Casey B. Mulligan
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
and NBER
c-mulligan@uchicago.edu

Barbara Schneider
Department of Sociology
University of Chicago
1126 East 59th Street
Chicago, IL 60637
schneidr@norcmail.uchicago.edu

Rustin Wolfe
300 E. 4th St. #717
St. Paul, MN 55101
rustw@dura.spc.
uchicago.edu

Do studies of time use interfere too much in the lives of the subjects? As a result are those who agree to participate a biased sample of the population? This has been claimed to be the case for the Michigan time-diary studies by Hochschild (1989, p. 273), Leete and Schor (1994), and others.[1] The purpose of this paper is to examine the characteristics of the Experience Sampling Method (ESM) adolescent sample from the Alfred P. Sloan Study of Youth and Social Development in order to detect and quantify some instances of nonresponse bias.

Created by Mihalyi Csikszentmihalyi and colleagues (Csikszentmihalyi and Csikszentmihalyi, 1988; Csikszentmihalyi and Larson, 1992), the ESM was primarily designed to examine how individuals spend their time, what they do, and what their subjective interpretations of their emotional states are during specific activities. Individuals are given beepers or programmable wrist watches that are randomly activated throughout the day. When beeped, the respondent fills out a self-report of what he or she is doing and how he or she feels at that moment.

The ESM has been criticized as being too burdensome and that there may be an inherent selection bias with the method if people who agree to participate in the study differ systematically from people who do not agree to participate (Zuzanek, 1999). It has also been suggested that individuals may underreport what they are doing because they do not wish to be interrupted. Such underreporting is assumed to occur more frequently during activities outside the home where respondents may be unwilling to answer the beep. To deal specifically with the questions of nonresponse bias and underreports of various activities, we compare Sloan data from over 1,000 adolescents who used the ESM with other adolescent data drawn from the Current Population Survey (CPS) and the National Education Longitudinal Study of 1988-94 (NELS:88-94). These national

[1]Defense of time use methodologies can be found in Juster and Stafford (1991) and Robinson and Godbey (1997). Gershuny et al (1986) reports some effects of length of subject study on response rate and the nature of the sample responding in studies of adult time use.

studies use sampling procedures specifically designed to represent the general adolescent population and their respective sampling parameter estimates are robust.

We focus on two dimensions of sample selection bias, representativeness of sample and nonresponse bias. The Sloan adolescent sample was not designed to randomly sample American teenagers. Individuals who use these data may therefore question how representative the sample is in comparison to studies designed to make inferences about the American teenage population. With regard to nonresponse bias, the ESM, because of its respondent burden, may result in certain activities being underreported. Our findings are useful not only for researchers interested in using the Sloan data, but for others who are considering analyzing other ESM data.

There are a variety of dimensions along which the Sloan ESM sample and other samples might be compared. Our focus is on time use, that is whether Sloan adolescents' time use, which is constructed from repeated measures, is similar to time use obtained through single point responses reported by adolescents in other studies. We confine our attention to those variables which (1) proxy for time use, such as watching television and hours working outside of school, or (2) are asked of the Sloan ESM sample and adolescents in other surveys, such as gender, age, number of siblings, performance in school, and days of the week and time of year working for pay.

We find the Sloan ESM sample to be representative of general populations in many, but not all dimensions. The sample is nearly representative in terms of teen employment rates, parental employment rates, a student's grade point average, and TV watching. Work hours are slightly undercounted in the study because of slightly higher nonresponse rates by teenagers working long hours.

The sample is less representative in terms of the time of week and gender; nonresponse is relatively common on school nights and (to a lesser extent) on weekends, and among boys. Sloan ESM is less than seasonally representative, with more observations in April, May, and October. However, the Sloan sample includes a significant number of observations for all nine of the academic months of the year and is therefore more seasonally representative than a number of other studies of adolescents. We use our time-of-week and gender results to construct a set of weights which analysts might use to estimate statistics for the general adolescent population.

I. Sloan Study Design

The Alfred P. Sloan study is a national longitudinal study that began with grade cohorts in sixth, eighth, tenth, and twelfth grades. Over a five year period, 1,221 students were followed in 12 sites throughout the U.S. The sample was drawn in three stages: localities, then schools within each locality, and finally students within each school. Localities were selected to satisfy the following criteria, variation in urbanicity, labor force composition, and race and ethnicity. Using 1990 U.S. census information, 15 potential sites were selected based on the degree to which their local economies were concentrated in manufacturing or service, as well as in their trend toward economic growth, stability, or decline over the past decade.

Once the sites were selected, local area educational superintendents were contacted. The superintendents were asked to identify those high schools that they considered the "most typical high school" in the district with respect to student demographic characteristics and college attendance rates. Based on their recommendations, high school principals were contacted. At that time, high school principals were asked to identify the elementary or middle schools that feed students into their respective high schools. Based on the responses of the willingness of superintendents and high school principals to participate in the study, twelve sites across the U.S. were obtained from the original list of 15. The twelve sites were matched so that several comparisons could be made among the school communities regarding the socioeconomic status and racial and ethnic diversity of the school population and school size. To ensure racial and ethnic diversity among some of the schools, higher numbers of middle class African Americans and Hispanics were oversampled relative to their proportions in the national population. This purposive sampling plan was undertaken since much of the career literature had excluded these populations.

The 12 sites included 33 schools: 20 middle schools and 13 high schools. To provide variation in high school programs, two specialized schools were included in the sample--a mathematics and science academy and the other a magnet language academy. The remaining 11 high schools had more traditional comprehensive curricular programs. A small honorarium of $250.00 for each year of participation was offered to each school.

For each elementary or middle school and each high school, two student samples were selected: ESM focal students and Questionnaire-only (Q) students.[2] The focal students were chosen

[2]Schneider and Stevenson (1999) refer to the ESM and Q students as "focal" and "cohort" students, respectively.

from school-prepared enrollment lists of grades 6, 8, 10, and 12. Using a stratified design at each school, student selections at each grade level were made so that they were proportionately representative of gender, race, ethnicity, and level of academic performance. Based on student records, teachers rated each of these students as academically successful, working at grade level, or having academic problems. At each school twenty-four students from each grade level were selected from lists prepared by the school using a random table of numbers.

The Q sample was selected using the same criteria as the focal sample. The Q sample was designed to provide more information about the school and peer networks for each of the focal panel grades. Each year in the field, new Q samples were drawn from the grade the focal students were in. If a grade enrolled no more than 150 students, the Q sample consisted of the entire grade. Otherwise a random sample of 150 students was chosen from the grade enrollment lists.

The Q students were administered most of the same instruments as the focal students making it possible to aggregate information from both sample groups. Combining the focal and Q samples over the five years of the study the total sample of Sloan study is over 8,000 students. Analyses of the focal and Q samples revealed no differences in demographic characteristics, attitudes toward school, educational expectations, occupational aspirations, and other key variables used in this study.

Data were collected from the focal students by three methods: (1) the Experience Sampling Method; (2) an in-depth interview (revised each year the adolescent was contacted); and (3) a battery of questionnaires. The questionnaires included the Teenage Life Questionnaire (a modification of instruments used in the National Education Longitudinal Study of l988-94), a Friends Sociometric Form, which provides information regarding adolescents' peer groups, and a questionnaire called the Career Orientation Scale, which measures job knowledge and occupational expectations. The Q students completed the questionnaires but were not interviewed and did not participate in the Experience Sampling Method.

Once the students graduated from high school only the ESM sample was followed. New interview forms and brief telephone interviews were conducted with the focal sample. Questions in these instruments focused on college and work experiences. Special forms were designed for young adults in college, in college and working, working but not in college, and not working and not in college. Through the five years of the study, 84% of the focal sample has been retained. Our study uses the first (1993) wave of ESM and Q students and the fifth (1996-7) wave of ESM students.

After completing a questionnaire pertaining to family characteristics, experiences in school, and plans for the future, the ESM students wore wrist-watches programmed to beep randomly eight times daily in intervals between 7:30am and 10:30pm on a schedule of eight days and 56 intervals.[3] Students completed a short questionnaire describing their activities and thoughts at the time of the beep. After the data were gathered, eight time slots were generated to generalize about the various time slots across sites, schools, and cohorts: 7:29-9:17am, 9:18-11:10am, 11:11am-1:03pm, 1:04-2:57pm, 2:58-4:49pm, 4:50-6:42pm, 6:43-8:35pm, and 8:36-10:14pm.

As explained in detail below, subjects did not respond to all beeps, but the number of beeps to which a subject responded and was engaging in a particular activity can be used to approximate a percentage of time engaged, and an absolute amount of time for the week. To do so, we first calculate the ratio *r* of beep responses while engaged in the activity to the total beep responses. Since beeping occurs approximately 15 "waking" hours each day (more precisely, 7:29am-10:14pm or 886 minutes) over a seven-day week, 105 waking hours (more precisely, 6202 minutes) per week are represented. So each percentage point of *r* corresponds to 1.05 weekly hours (more precisely, 62.02 minutes). For example, we find 10.6% of beep response to occur while the subject watched TV (as a primary activity, see section IX), so we estimate 11 weekly waking hours (657 weekly waking minutes) of watching TV.

II. Our Comparison Groups

As comparison groups, we use the Census Bureau's May 1993 Current Population Surveys (CPS) and the Education Department's 1988-94 National Education Longitudinal Study (NELS:88-94). The CPS has two advantages. First, it is designed exclusively to obtain national estimates population demographics and labor force activity (Census Bureau 2000, p. 3-1). Unlike the decennial Census which relies on a great many citizens completing and returning the questionnaire under no direct Bureau supervision, the CPS respondents are statistically sampled, and then located and

[3]An eight-day schedule was used for all subjects, although the day of the week on which the schedule began varies by site, school, and cohort. While the beep cycle starting day and ending days varied, the studied was designed to begin between 11:11am-1:03pm of the first day and to end between 9:18am-11:10am of the eighth day – a total of seven complete days and 56 beeps.

questioned by trained interviewers (Census Bureau 2000, pp. 7-1f).[4] Teenage employment may be seasonal, so another advantage of the CPS is that we can use its monthly surveys to examine and, if necessary, correct for the effects of seasonality in our data.

The main disadvantage of the CPS is that is does not ask many of the questions of interest to users of the Sloan Study (such as the employment and other uses of time of those age 14 and under), so we supplement with NELS:88-94 comparisons. NELS:88-94 is a nationally representative sample of adolescents that began in 1988 when 25,000 students were enrolled in public and private high schools in the U.S. The students were resurveyed in 1990 when most were in tenth grade and in 1992 when most are in the 12th grade (Ingels, Scott, and Taylor, 1997). The data collected included information from students, parents, teachers, and school administrators. In addition to basic demographic and family information, NELS:88-94 includes variables measuring performance in school, educational aspirations, experiences in school and experiences at work.

The purpose of these comparison groups is to isolate the two dimensions of sample selection bias: Sloan design and nonresponse bias. As is evident below, we attempt to isolate the first by comparing the Sloan "Q-only+ESM" sample of teens responding to a questionnaire, but not necessarily to the ESM, with CPS and NELS samples – under the assumption that CPS and/or NELS respondents adequately represent the teen population. We attempt to isolate the second dimension by comparing the "Q-only+ESM" sample with ESM samples.

III. Day-of-Week in the Sloan Study

As described above, ESM beeps occurred at regular intervals during the waking hours for a calendar week. ESM students are somewhat less likely to respond to a beep if it is on the weekend or on a school night. This pattern can be seen rather simply is our Table 1, which tabulates beeps according to four partitions of the week: "school time" (7:29am-2:57pm Mon-Fri), "after school weekday" (2:57pm-6:42pm Mon-Thu), "school night" (6:42pm-10:14pm Sun-Thu), and "weekend"

[4]See Hogan and Robinson (1993) for a discussion of how the Census Bureau uses statistical sampling like the CPS to estimated undercounting in the decennial census, and how it appears that black and other minorities are undercounted in the census.

(2:57pm-10:14pm Fri, all day Saturday, or 7:29am-10:15pm Sunday).[5] Since these four partitions of the week are of different duration, Table 1 first column displays what the allocation of responses across partitions would be if the responses were truly random. The second column shows the allocation of actual responses.

Table 1: Time-of-Week Distribution of ESM Wave 1 beeps
(of those reporting at least 15 beeps)

Time-of-Week	random percentage	% ESM beeps	response rate (%)
school time	35.7	42.1	68.6
after school weekday	17.9	19.0	61.8
school night	17.9	16.0	52.0
weekend	28.6	22.9	46.6
TOTAL	100.0	100.0	58.1
beeps		26603	

Since, the third column shows the ratio of the first two columns we see, for example, from the Table how ESM responses are 18% too frequent during school time (as compared to the hypothetical with 41.8% (=100-58.2) nonresponse allocated randomly across partitions of the week according the each partition's duration), and 11% too rare on a school night. The time-of-week response rate differential is quantitatively and statistically significant, with response rates almost twice as high during school time as on a school night. Although response rates are related to gender (see

[5]Our partitions of the week are irregularly timed in order to conform to the Sloan study's beeping schedule. That schedule was constructed first by dividing the day into two hour intervals, and then randomly generating 7 times for each interval for the entire study. Of course, the earliest of the seven beeps generated was not precisely at the beginning of the interval, and the latest was not precisely at the end (for example, the earliest beep was 7:34am even though the interval for random generation began at 7:30). Our time of week partitions conform to the earliest scheduled time in an interval, minus 5 minutes for potential desynchronization of study watches (for example, our partitions begin the day at 7:29am = 7:34am - 5 minutes).

Table 4), the time-of-week differential response is not.

Although participation in the sample of students reporting at least 15 beeps declines with age (see our Table 3 below), we do not find a systematic relationship between age and beep response rate, even when interacted with time of the week. On the first point, the overall beep response rate by grade ranges from 56% for seniors to 59% for 6th graders. As one example of the second point, we notice that the beep response rate on school nights ranges from a low of 36% for 6th graders to a high of 42% for tenth graders.

IV. Seasonality in the Sloan Study

The Sloan Study is designed for the analysis of teenagers in working families. Hence, only adolescents who are in school are studied and, by design, the resulting data on time use is not representative of the calendar year. Furthermore, each wave of the Sloan study was conducted during only two or three calendar months. We therefore advise users interested in a seasonally representative sample to pool all waves of the Sloan study, which we do for the purposes of calculating Table 2.

As we see in Table 2, most of the Sloan observations are in April, May, or October. November through March are also represented, but summer months are basically unobserved since the Sloan study was designed to observe adolescents in school.[6,7] Although there are significant differences between the ESM sample and a seasonally random sample, the seasonal differences between the ESM and ESM 15 samples are quantitatively and statistically insignificant.

It is important to note that in other surveys such as NELS:88-92, data collection occurs during several months, primarily February through April. We would expect that teenagers reports of employed work in NELS:88-92 would be subject to similar seasonal variations. Since most teenagers frequently change jobs, working different number of hours at different times, we could assume that seasonal variations in employed work among teens is relatively random. Exceptions would occur during the winter holiday season and over the summer as more employment opportunities exist for

[6]Wave 1, 2, and 3 observations are mainly for April and May of 1993, October 1994-January 1995, and October 1996-February 1997, respectively.

[7]It should be noted that those few (6%) observations in June or September are for students who are in school during those months.

teenagers and they tend to be responsive to these labor market opportunities (Protecting Youth at Work, l999).

Table 2: Seasonal Distribution of ESM Waves 1, 2, and 3 respondents

Month	% ESM students	% ESM w/ num beep ≥ 15
January	6	4
February	7	8
March	7	9
April	19	20
May	20	21
June	5	6
July	0	0
August	0	0
September	1	1
October	17	15
November	9	9
December	8	7
TOTAL	100	100
respondents	2078	1479

Notes: (1) Month for Q-only students is estimated according to the month ESM students in the same school were observe

(2) ESM Students observed in during two months are tabulated according to the month of their last observation.

V. Basic Demographics in the Sloan Study

Table 3 displays the age distribution of adolescents in the CPS and two ESM samples. The overall ESM sample is representative of the age distribution of those in school, and reflects the 11% High School dropout rate seen in the CPS and other surveys. The propensity of an ESM student to provide at least 15 beeps declines with age. The age differences between the ESM and ESM 15 samples are quantitatively insignificant, although they are statistically significant at the 95 confidence level. The age differences between the ESM 15 and CPS samples are both quantitatively and statistically insignificant.

Table 3: Age Distribution of ESM Wave 1 respondents

Grade	May 1993 CPS	NELS	Q-only + ESM	% ESM students	% ESM w/ numbeep ≥ 15
Grade 6	27	0	16	27	29
Grade 8	26	36	24	28	28
Grade 10	25	33	36	24	24
Grade 12	22	31	24	21	19
TOTAL	100	100	100	100	100
respondents	8072	54974	4113	1109	865

Notes: (1) Procedure for assigning CPS grades: it is assumed that all children 15 and under are enrolled in school in the week prior to the May interview. Those age 11, 13, and 15 are assumed to be in the 6^{th}, 8^{th}, and 10^{th} grades, respectively. Those age 17 and enrolled in high school assumed to be in the 12^{th} grade.

(2) NELS Age distribution computed by pooling the 1988, 1990, and 1992 waves (which all interview the same cohort). NELS "respondents" therefore double and triple count some individuals.

As we look across the columns of Table 3, we begin to see some of the different contributions

of two dimensions of sample selection bias – the Sloan sample design and its unique survey instrument (ESM). Consider, for example, the Grade 12 row. We do not see a big difference between CPS, Q-only, and NELS (adjusting for noninterview of 6th graders) which suggests that the Sloan study was not designed in such a way to under- or over-represent high school seniors relative to those in the 6th, 8th, and 10th grades. However, a comparison of the last three columns suggests that the ESM does contribute somewhat to under-representation of seniors because the ESM sample has a relatively small percentage of seniors, and the ESM sample of those responding at least 15 times has an even smaller percentage. In other words, seniors have a lower response rate to the ESM, and will be somewhat under-counted in an unweighted beep-level analysis.

Most of our analysis presumes that the "Q-only + ESM" sample adequately represent the population sampled by Sloan for the ESM study, and differences with the ESM derive from nonresponse. Based on the Sloan study design (see our Section I above), we believe this to be true in most dimensions – but not when it comes to comparing the incidence of six graders in the sample. Remember that, at each school site and for each grade (6, 8, 10, 12), a focal group of students was targeted for administering the ESM and then (up to a sample of 150) *all* other students in that grade at that site was targeted for administering the questionnaire. Since sixth graders typically school at smaller sites (eg., the sixth grade site will often be small neighborhood sites rather than large high schools or junior highs schooling the grade for the entire school district), this design implies: (a) a lesser incidence of sixth graders in the "Q-only + ESM" sample than in a "representative" sample such as the CPS, and (b) a lesser incidence of sixth graders in the "Q-only + ESM" sample than in the ESM focal sample – even if ESM response rates were uncorrelated with grade. Hence, with regards to Table 3, the "Q-only + ESM" sample is not particularly helpful for isolating the two dimensions of sample selection bias.

Table 4 shows that girls are more likely to participate in the Sloan study, and more likely to provide 15 or more responses during the week. Our tabulation of the CPS sample strongly suggests that only a small minority of the differential participation is due to there being more girls in the school population – 48 or 49% of the school population is male while only 41% of those responding with 15 or more beeps is male.

Table 4: Age and Gender Distribution of ESM Wave 1 respondents

		Grade					
sample	gender	6	8	10	12	all	respondents
% CPS sampled in grade that are:	male	49	49	49	48	49	3896
	female	51	51	51	52	51	4176
	either	100	100	100	100	100	8072
% ESM students in grade that are:	male	45	47	44	43	45	498
	female	55	53	56	57	55	611
	either	100	100	100	100	100	1109
% ESM students w/ numbeep ≥ 15 in grade that are:	male	40	44	41	39	41	357
	female	60	56	59	61	59	508
	either	100	100	100	100	100	865

Notes: Procedure for assigning CPS grades: it is assumed that all children 15 and under are enrolled in school in the week prior to the May interview. Those age 11, 13, and 15 are assumed to be in the 6^{th}, 8^{th}, and 10^{th} grades, respectively. Those age 17 and enrolled in high school assumed to be in the 12^{th} grade.

VI. Number of Siblings

We find that the Sloan study slightly over-represent adolescents with multiple siblings.

Table 5: Number of Siblings of ESM Wave 1 respondents

Month	May 1993 CPS	Q-only + ESM	% ESM students	% ESM w/ numbeep ≥ 15
none	25	24	23	22
one	39	29	29	30
two	22	24	25	24
three	9	13	12	13
four+	5	11	11	11
TOTAL	100	100	100	100
respondents	7591	4113	1109	865

Table 5 also illustrates how we believe the CPS, "Q-only + ESM", and ESM samples can isolate the two dimensions of selection bias. A comparison of the first two columns suggests that the Sloan study somewhat oversamples adolescents with one sibling, while a comparison of the last three columns suggests that nonresponse to the ESM is basically uncorrelated with number of siblings.

VII. Teenage Employment

When it comes to using the ESM to measure time use, and work time in particular, there are three issues that must be addressed by our study. First, how well do those Sloan study students reporting at least 15 beeps represent the teenage population in terms of work histories, or current work status? Second, are Sloan study work-related questionnaire items comparable to work-related questionnaire items from the NELS or CPS? Third, how do ESM estimates of time use compare with estimated derived from questionnaire responses?

VII.A. Work-related Questions in the CPS, NELS, and Sloan study

In order to deal with the first and third points, we need to address the second. In particular, the Sloan study's work-related questions have important differences with those in the CPS. The Sloan study asks "Are you currently employed (have a paying job) or have you ever been employed?"

to which valid responses are “never,” “not employed now but was employed during this school year,” “not employed this school year but was employed last summer,” “was employed prior to last summer,” or “currently employed.” Note in particular that “currently” is rather open-ended – does it refer to the day of the interview, the week of the interview, the month of the interview, or the semester of the interview? Also, does baby-sitting, yard work, or work at the family business count? We expect these distinctions to be more important for teens than for adults, since the former are less attached to the labor force, and their time spent in schooling makes irregular, intermittent and/or informal employment relatively more attractive.

However we answer these questions, we see in the last three columns of Table 6 how there are only minor differences across Sloan samples in the fraction of students “currently” employed. The slight difference between the “Q-only+ESM” and ESM15 samples in the employment rate of high school seniors suggests that the schools and/or regions targeted by the Sloan Study slightly overrepresent the population of working teenagers, although “Q-only+ESM” employment rates are higher for those in grades 6, 8, and 10.

The closest question in the NELS88-94 study is “What is your job situation?”. We see in Table 7 that the fractions “currently working” are pretty similar in the NELS and Sloan samples.

Table 6: Currently Working Status of ESM Wave 1 students – Questionnaires

	May 1993 CPS	NELS	Q-only + ESM	ESM students	ESM w/ numbeep ≥ 15
percent working	NA	NA	38.0 [1756]	38.9 [386]	37.9 [319]
average hours among those working	NA	NA	17.8 [654]	17.2 [148]	16.4 [119]
percent working seniors	39.5 [1927]	51.7 [16070]	52.5 [708]	54.7 [170]	54.3 [138]
average hours among working seniors	17.4 [778]	16.7 [7601]	19.0 [365]	18.8 [93]	17.7 [75]

Notes: (1) Procedure for identifying CPS seniors: those age 17 or 18 or 19, enrolled in high school, have completed the 11th grade, but do not have a HS diploma.

(2) CPS observations are weighted using the household head's CPS weight

(3) number of observations reported in brackets

(4) In the Sloan study (Q-only and ESM samples), working is indicated by a "currently employed" response to the question "What is your job situation?".

(5) In the Sloan study, hours working on current job are reported in 0-10, 11-20, 21-30, 31-40, and 41+ hour intervals. We used the Sloan distribution of responses across these intervals, and CPS interval averages for seniors (7.0, 16.8, 25.8, 36.5, and 46, respectively), to compute a Sloan average hours.

(6) NELS statistics are weighted according to the NELS variable F2QWT, which weights the second follow-up sample to represent the 1992 U.S. population of 12th graders.

The CPS questions about employment status are more specific, and we use the Census Bureau's concept of "currently employed and working" derived from those questions.[8] In particular, "currently employed and working" refers to those who worked for pay some time during the survey week, plus those working 15 hours or more as unpaid family workers during the survey week. We

[8]ie, those who "during the survey week, do any work at all as paid employees or in their own business or profession, or on their own farm, or who work 15 hours or more as unpaid workers on a farm in a business operated by a member of the family." (Census Bureau 1995, pp. 22-3)

see in Table 7 that, according to the various questionnaires, the CPS fractions of seniors currently working are substantially lower than in the NELS and Sloan samples. Given that NELS and CPS represent well the teen population, this difference appears to be attributable to the survey question rather than populations sampled (see also Committee 1998 pp. 40f).

Sloan, NELS, and CPS ask about weekly hours usually worked on the current or most recent jobs, and we report the average for those "currently employed" in the second and fourth rows of Table 7. Notice how those 17% of the ESM reporting less than 15 beeps tend to work longer hours if they are employed. The average hours differences between the ESM and ESM15 samples are statistically significant and of some quantitative significance, although neither the ESM nor the ESM15 is statistically significantly different from the Q-only sample.

As discussed above, the Sloan employment question can be used to measure whether a respondent has ever worked in his/her lifetime. NELS responses to the "What is your job situation?" question can be used to determine whether a respondent ever had a job in his/her lifetime. Table 7 suggests that ESM students reporting at least 15 beeps represent well the overall population in terms of propensity to work or have worked; the differences between the fraction working in that sample is quantitatively and statistically insignificantly different from that for the ESM and Q-only samples.

Table 7: Lifetime Ever Worked Status of ESM Wave 1 students

	NELS	Q-only + ESM	ESM students	ESM w/ numbeep ≥ 15
percent ever worked	NA	67.9	70.5	70.2
		[1756]	[386]	[319]
percent ever worked	85.8	84.5	87.1	85.5
seniors	[16070]	[708]	[170]	[138]

Notes: (1) number of observations reported in brackets

VII.B. ESM as a Measure of Work Time

The ESM can be used to measure employment and hours, and in a way that is comparable to the CPS's "survey" week definition of "currently employed and working". To measure employment

we take, in the sample of students responding to 15 or more beeps during the week, the fraction of them reporting at least one beep in the workplace, as shown in Table 8. We find an "employment rate" for seniors that is both similar to the CPS questionnaire-based estimate and substantially different from the fraction of those responding affirmatively to the Sloan study's rather open-ended "currently" employed question.

Table 8: Working Beeps of ESM Wave 1 students

survey instrument:	questionnaire	beeps	
		unweighted	weighted
Sample:	May 1993 CPS	ESM w/ numbeep ≥ 15	ESM w/ numbeep ≥ 15
percent working	NA	15.7 [865]	14.7 [865]
average hours among those working	NA	11.3 [137]	11.8 [137]
percent working seniors	39.5 [1927]	39.3 [168]	36.7 [168]
average hours among working seniors	17.4 [778]	15.2 [67]	15.9 [67]

Notes: (1) Procedure for identifying CPS seniors: those age 17 or 18 or 19, enrolled in high school, have completed the 11th grade, but do not have a HS diploma.
(2) number of observations reported in brackets
(3) a "working beep" is one that occurred while the subject was at his workplace (eg., even if during breaktime)

The reporting of *at least one* beep in the workplace is the obvious measure of employment, but beeps can be aggregated to obtain an estimate of the number of hours worked during the survey week. To do so, we first calculate the ratio *r* of beep responses at work to the total beep responses.

Since beeping occurs approximately 886 minutes per day over a seven-day week, or 6202 minutes per week. So each percentage point of *r* corresponds to weekly 62.02 minutes. For example, for those responds reporting at least one beep at work, we find 11.0% of beep responses to occur while the subject was at work, so we estimate 11.3 weekly waking hours at work for those who worked at all.

We know how ESM response rates vary by time of week, and by gender, so we might reweight ESM responses by the inverse of the response rate for that gender/time of week. In particular, we expect (and find) much less work during school time – and school time has the best response rate – so we expect a tome-of-week-weighted estimate of work time to be higher.[9]

The weighted and unweighted hours estimates can be compared with the hours estimates from made by respondents on their questionnaires. Our ESM estimates of the employment rate, and weighted ESM estimates of hours worked, are *very* similar to CPS questionnaire-based estimates. To the extent there are differences, they might be interpreted in three ways:

(1) Reported beeps are imperfectly representative of adolescent time use (eg., ESM nonresponse is especially high when the respondent is at work)

(2) questionnaires estimates of the length of the workweek are imperfect

(3) CPS and Sloan sample different populations

The second interpretation has been made by authors of time diary studies of the workweek. In particular, it is found that adults with shorter workweeks overestimate their workweek the most. If this reported bias carried over to teenagers, we would expect questionnaires to overestimate teenage work, since their workweeks are short relative to that of an average adult. However, there are a few reasons to suspect questionnaire biases to be different for teenagers than for adults. First, teenagers are typically "clockpunching" hourly employees, and the process of punching the clock permit them a better estimate of work hours than for (typically adult) salaried employees. Second, teen work schedules are much less regular, which makes it less likely that a teenager would accurately estimate his work hours for any given week.

Given the similarity of ESM and CPS estimates, our data do not suggest that the *CPS*

[9]The ESM suggests that girls tend to work more, so the gender dimensions of the weights tends to reduce estimated employment rates and work time, although this effect tends to be smaller in magnitude than the time-of-week dimension of the weights.

questionnaire produces systematically biased estimates of teen *time at the work place*.[10] However, our data do suggest: (a) that the Sloan and NELS questionnaires substantially overestimate teen time at the workplace and (b) that teen time at the workplace is not the same as teen time worked. The first suggestion is made by our Tables 7 and 9, which show how teen employment measured by Sloan and NELS questionnaires is substantially higher than employment measured by the ESM or the CPS questionnaire.

The second suggestion is made by looking at what teens were doing when at work, as shown in Table 10. When beeped at work, teens reported working 80% of the time, and the other 20% of the time reported doing homework, talking with friends, playing games, watching TV, listening to music/radio, doing a hobby, personal care, or smoking. For some applications, these other activities may be considered something other than "work" – even though done at work – and the ESM shows that they are nontrivial and offers researchers some quantitative indicators of those activities.

[10]Gershuny et al (1986) have a similar finding in their British study of adult time use – they find similar average hours worked in a time diary sample and in a more standard employment-questionnaire sample. Perhaps surprisingly, their point estimates suggest that those who work long hours are *more* likely to respond in a diary study than in an employment-questionnaire study.

Table 9: Activities at Work,
ESM Wave 1 students reporting at least 15 beeps
(percent of beeps at work)

	unweighted	weighted
working as primary activity	73.4	73.8
working as secondary activity	6.4	5.9
working neither as primary nor secondary activity:		
homework	0.6	0.6
talking with friends, in person	6.8	6.7
talking with friends, other	1.4	1.3
playing games	1.6	2.0
watching TV	3.9	4.0
listening to music/radio	0.2	0.3
doing a hobby	1.6	1.9
personal care	3.9	4.0
smoking	0.2	0.2

Notes: The following activities are coded as "work" when done at the workplace: "thinking", "standing" "walking", "waiting", "driving", "nothing", "missing the beep" or "this study" (!)

VII. Parental Employment and Occupation

We see in Table 6 that the work status of parents of ESM students reporting at least 15 beeps is representative of the overall population. There is a slight, but statistically insignificant, tendency for the sample to overrepresent students with father only working and underrepresent students with both parents working relative to the ESM and Q-only samples.

Table 10: Work Status of ESM Wave 1 parents

Month	May 1993 CPS	NELS	Q-only + ESM	% ESM students	% ESM w/ numbeep ≥ 15
neither parent works	6.7	3.7	2.1	2.0	2.0
only father	27.4	10.4	13.6	15.1	15.7
only mother	6.8	9.0	4.5	4.0	4.0
both work	59.1	76.9	79.9	78.9	78.3
TOTAL	100	100	100	100	100
respondents	5738	19379	2851	697	599

All of the Sloan samples have a higher fraction of two parent working families as compared with the CPS. This difference is mainly due to the different questionnaire items in the two studies, but we believe that some of the difference is real because the Sloan study was designed to study "working families."

IX. TV Watching

For 10.6% of the ESM responses (by those responding to at least 15 beeps during the week), "watching television" was reported to be the main activity, and a secondary activity for another 4.1%. Since the ESM samples the 15 hours of the day 7:30am-10:30pm, and assuming little TV watching between 10:15pm and 7:28am, 10.6% of beeps is 96 minutes per day (131 minutes including TV as a secondary activity), or 11 hours per week (15 hours per week). Table 11 reports these "unweighted" results in the 2nd-to-last row to facilitate comparison with other studies.

Table 11: Teen TV watching time by Study and Survey Instrument

study	instrument/year	age group	TV hrs/wk primary	TV hrs/wk primary + secondary
Robinson & Godbey (1997)	time diary/1985	12-17	17	
Gallup (1993)	question kids/1993	13-17		11
NCES (1993)	question kids/1990	15		17
Annenberg (1996)	question parents/1996	2-17		18
Annenberg (1997)	question parents/1997	12-17		15
Gunter and McAleer (1997)	TV meter/1993	12-15		20
Sloan study - unweighted	**ESM/1993**	**12-18**	**11.0**	**14.7**
Sloan study - weighted	**ESM/1993**	**12-18**	**12.1**	**16.5**

We know that how ESM response rates vary by time of week, and by gender, so we might reweight ESM responses by the inverse of the response rate for the corresponding gender-time of week cell. In particular, we expect (and find) much less TV watching during school time – and school time has the best response rate – so we expect a tome-of-week-weighted estimate of TV to be higher. Table 11 verifies this – weighting adds an hour or two per week.[11]

Robinson and Godbey (1997, p. 209), using a diary method for measuring time use, find 17 hours of TV watching (as a primary activity) per week for teenagers in 1985.[12] Gallup's (1993) questionnaire-based study suggests that those aged 13-17 watch TV about 10 or 11 hours per week,[13] while NCES's (1993) questionnaire-based study of 1990 high school sophomores suggests 17 hours

[11]Because TV watching varies less by gender than time of day in our sample, most of the effect of weighting is due to the time-of-week dimension of the weights.

[12]They sample all of the months of the year.

[13]Gallup (1993) tabulate respondents aged 13-15 and 16-17 by intervals of TV watching hours for the day before the interview: none, 0-1 hours, 1-2 hours, and more than 2 hours. Assuming interval averages of 0, 0.5, 1.5, and 4, respectively, Gallup's (1993) tables suggest 1.5 hours per day or 10-11 hours per week.

per week. Gallup (1993) and NCES (1993) questioned teens, while Annenberg (1996) questioned parents of children (including preteen children) about their children's TV watching time and found about 18 hours per week. Using meters installed in television sets, Gunter and McAleer (1997)[14] found children aged 12-15 watching almost 20 hours of TV per week in 1993.

The ESM is somewhat unique in distinguishing TV watching as a primary rather than a secondary activity. It seems that, other than the time diary study, the studies mentioned above would include TV watching as a secondary activity (eg., while doing homework), so 16 hours per week may be the comparable unweighted estimate. Hence, since a variety of other measurement methods suggest 17 hours or more per week for teens in the 1990's, it seems that the ESM offers a close, but slightly underestimated, measure of TV watching time. Perhaps the underestimate is not surprising, since our calculations assume no TV watching after 10:15pm.

X. Performance in School

Both NELS and Sloan questionnaires include questions posed to 10th graders about grades in four subjects (english, math, science, and social studies), which we aggregate for each respondent in both studies to compute a grade point average on a four point scale.[15] Table 12 displays the grade point averages for 10th grade students in the NELS and three Sloan study samples. They are fairly similar for all of the samples, although the small gap between NELS and Sloan is statistically significant. Judging from the 3.07 GPA of the Q-only sample, the main difference appears to be the schools sampled by the Sloan Study (compare 3.07 with 2.89 for NELS respondents), rather than nonresponse within the Sloan study. The average GPA for the 14% of Sloan ESM 10th graders not responding with at least 15 beeps is 2.7, for a GPA difference of only 0.08 between the ESM and ESM15 samples.

[14]Their data is from the Broadcasters' Audience Research Board in the U.K., and the "metering" system involves placing meters in sample households' TV sets which record when the TV is turned on, for how long, etc., and requires that each family member identify themselves with a remote control each time they use the TV, and this gets recorded.

[15]Grading questions asked of 8th and 12th graders are tougher to compare across the Sloan and NELS studies.

Table 12: School Performance of 10th Graders

Month	NELS	Q-only + ESM	% ESM students	% ESM w/ numbeep ≥ 15
avg GPA	2.89	3.07	3.12	3.20
respondents	13735	999	200	171

XI. Weights for Sloan Study Users

Although unweighted ESM estimates of time use are fairly close to estimates from other studies, ESM nonresponse is quantitatively significant in a couple of dimensions: time-of-week and gender. Estimates that better characterize the wider adolescent population, and more closely match estimates from other studies, can be obtained by weighting beeps according to their time-of-week and the gender of their respondent. Table 13 reports the weights we used in our analysis, expressed in proportion to the inverse of the probability that a beep would be included in the Wave 1 sample of beeps reported by those responding at least 15 times.

Table 13: Time-of-week-and-Gender Weights for
ESM Wave 1 beeps
(in sample of those reporting at least 15 beeps)

	gender		
time of week	male	female	either
school time	0.116	0.088	0.204
after school weekday	0.129	0.098	0.227
school night	0.153	0.116	0.269
weekend	0.171	0.129	0.300
any time	0.570	0.430	1

Note: each of the eight weights in the interior is the product of its time-of-week weight (from the last column) with its gender weight (from the last row).

The probability that a male (or female) respondent would be included in the sample is inferred by comparing the top (CPS) and bottom (ESM15) panels of Table 4. The probability that a beep at a particular time of week would be included in the sample is inferred from Table 1.

XII. Conclusions: Tradeoffs between ESM and Surveys

In selecting a particular method to measure time use, a researcher is confronted with several decisions. How can the study be economically administered? Can responses be compiled from a sufficiently representative sample? Can responses be expected to be accurate, and interpreted by subjects in as interpreted by study designers? Surveys such as the CPS and NELS:88-94 are practically useful in that they can be administered in a single session. In contrast, the ESM is certainly more difficult to administer, since participants are required to fill out response forms several times a day over an extended period of time. Selecting a population of subjects willing to complete the ESM would, we suspect, introduce some respondent selection bias. And, in fact, we find that girls are over-represented in the sample and among those who fill out the ESM. Older students are less

likely to respond to beeps, although we did not find a systematic relationship between age and beep response rate. Students who participated in the ESM are more likely to have more multiple siblings than national samples. Yet, with respect to specific characteristics of parents' employment, the work status of parents of ESM students appears to be representative of the overall population of households with adolescent children.

But as some have suggested, people who agree to complete the ESM may be more organized and diligent. We find that students who completed the ESM have slightly higher grade point averages than students in the NELS sample, with the highest grades being reported by those students who completed 15 or more beeps. These higher rates may be confounded by the fact that there are more females in the ESM sample, and girls tend to have higher grades than boys in elementary and high school.

The other more problematic issue regarding the ESM is the response rates by activity and time of week. We found that after school and weekend beeps are underreported. This problem can be handled through weighting procedures and we have shown how it is possible to weight the sample adjusting for nonresponse by time of week and for the overrepresentation of females. What is perhaps most surprising is that even though the ESM tends to have lower response rates after school and on the weekends, when estimating the percent of adolescents who have worked, the results from the ESM are nearly identical with national samples. These results suggest that ESM responses for reporting on activities outside the household and outside of school appear not to be as spurious as some have assumed that they may be. This comparability is also achieved when comparing CPS and ESM estimates of the average hours worked by high school seniors. However, if we weight the ESM sample by differential response patterns, we find that the percent who are working remains consistent with national samples but the average hours worked by seniors is slightly lower than the CPS.

ESM delivers a richer data set with repeated high frequency information on respondent time use. The study relies very little on subject recall, because responses are recorded at, and about, the time of the beep, and the time of the activity. This also mitigates the ambiguity present in one-time survey questions. For example, a one-time survey may ask “Are you currently employed?” Does that mean at the moment of the survey? The day of the survey? The month of the survey? In recent memory? There much less ambiguity when responding to the ESM – namely, was the respondent working or at work at the moment the beep occured? Employment for the day can be determined

by looking at all of the beeps for the day, for the week by looking at all of the beeps for the week, etc. Hence, it is not surprising that ESM measures of teen employment and hours are closer to those calculated based on the Census Bureau's one-time survey questions carefully designed (by trial and error, over a couple of decades of surveying) to measure weekly employment and hours than are those calculated based on less carefully designed and refined one-time survey questions such as "Are you currently employed?"[16]

Overall these results suggest that while there are sampling bias issues with the ESM, they are not terribly significant (at least for teen subjects) and can be attenuated through weighting. One-time survey questions about time use run higher risks that subjects will interpret them different than do other subjects and differently than do survey designers. As a result, the ESM offers more precise, or at least more robust and more easily interpreted, measures of time use, although perhaps of a somewhat selective sample.

XIII. References

Annenberg Public Policy Center. *Television in the Home 1996*. Philadelphia: University of Pennsylvania, 1996.

Committee on the Health and Safety Implications of Child Labor. *Protecting Youth at Work*. Washington, DC: National Academy Press, 1998.

Csikszentmihalyi, Mihaly and Isabella Selega Csikszentmihalyi. *Optimal Experience: Studies of Flow in Consciousness*. Cambridge: New York: Cambridge University Press, 1988.

Csikszentmihalyi, Mihaly and Reed Larson. *Being Adolescent: Conflict and Growth in the Teenage Years*. New York: Basic Books, 1984.

Gallup International Institute. *America's Youth in the 1990's*. Princeton, NJ: Gallup International Institute, 1993.

Gershuny, J. et al. "Time Budgets: Preliminary Analysis of a National Survey." *Quarterly Journal of Social Affairs*. 2, 1986.

Gunter, Barrie and Jill McAleer. *Children and Television*. 2nd ed.. New York: Routledge, 1997.

Hochschild, A. *The Second Shift: Working Parents and the Revolution at Home*. New York: Viking,

[16]Another advantage of ESM, not explored in this paper, is that it offers high frequency measures of subjective experiences.

1989.

Hogan, Howard and Gregg Robinson. "What the Census Bureau's Coverage Evaluation Programs Tell us about Differential Undercount." Washington, DC: Census Bureau, May 1993.

Juster, F. Thomas and Frank P. Stafford. "The Allocation of Time; Empirical Findings, Behavioral Models, and Problems of Measurement." *Journal of Economic Literature*. 29(2), June 1991; 471-522.

Leete, Laura and Juliet B. Schor. "Assessing the Time-Squeeze Hypothesis: Hours Worked in the United States, 1969-89." *Industrial Relations*. 33(1), January 1994: 25-43.

Robinson, John P. and Geoffrey Godbey. *Time for Life*. University Park, Pennsylvania: Pennsylvania State University Press, 1997.

Schneider, Barbara L. and David Stevenson. *The Ambitious Generation*. New Haven, CT: Yale University Press, 1999.

U.S. Bureau of the Census. *Current Population Survey, May 1993*. Ann Arbor, MI: ICPSR Study #6407, October 1995.

U.S. Bureau of the Census. *Current Population Survey – Design and Methodology*. Washington, DC: Census Bureau Technical Paper 63, March 2000.

U.S. Department of Education, National Center for Education Statistics. America's High School Sophomores: A Ten Year Comparison 1980-1990." Washington, DC: National Center for Education Statistics, 1993.

Zuzanek, Jiri. "Experience Sampling Method: Current and Potential Research Applications." paper prepared for the Time-Use Measurement and Research Workshop of the Committee on National Statistics organized by the National Research Council. Washington, D.C. May 27-29, 1999.