

Paper prepared for the "Conference on the Event History Calendar Method", organized by the University of Michigan's Panel Study of Income Dynamics (PSID) and the U.S. Census Bureau, December 5-6, 2007, Washington, D.C.

DRAFT ONLY – Do not cite or circulate

Event History Data: Lessons from the NLSY97

Charles Pierret (U.S. Bureau of Labor Statistics)*
Alison Aughinbaugh (U.S. Bureau of Labor Statistics)
A. Rupa Datta (National Opinion Research Center)
Tricia Gladden (U.S. Bureau of Labor Statistics)

November 2007

* pierret_c@bls.gov, 2 Massachusetts Ave. NE, Room 4945, Washington DC 20212, 202.691.7519. The views expressed are those of the authors and do not reflect the policies of the BLS or the view of other BLS staff members. We thank Rosella Gardecki for answering questions about the data.

The National Longitudinal Survey of Youth 1997 cohort (NLSY97) is a comprehensive nationally-representative social science survey sponsored by the U.S. Bureau of Labor Statistics. The mandate of the National Longitudinal Surveys program is to study long-term employment processes, but survey management has taken an expansive view of this mandate. As such, survey content includes not only employment activities per se, but also activities that can affect the ability to obtain and perform a job. Beyond comprehensive sections on education, employment, and training, the NLSY97 questionnaire includes extensive sections on marriage, fertility, household composition, income and assets, participation in government assistance programs, health, criminal behavior, and alcohol, tobacco, and drug use.

The basic unifying design concept of the NLSY97 questionnaire is event history interviewing. The instrument attempts to collect important events in the respondent's life that have occurred since her previous interview as well as to fully enumerate dates on which these events happened. Knowing these dates, the researcher can sequence events and know, for instance, whether the respondent changed jobs before or after the birth of her child. This helps establish causation, because only an event that precedes another can be said to have caused the latter event. The backbone of the NLSY97 data set is a series of complete event histories on education, training, employment, marriage and cohabitation, fertility, and participation in government assistance programs.

When collecting event history data, survey designers must tackle a number of issues concerning the design of the questionnaire and the dissemination of the data. This paper discusses how the NLSY97 design team addressed some of these issues and what resulted from these decisions. We start with some more background about the NLSY97 and then examine five issues that confronted NLSY97 designers and confront survey designers in all longitudinal social

science surveys. We hope the lessons we have learned will help others to address these issues in other surveys.

NLSY97 Background

The NLSY97 sample is comprised of 8,984 youths born between 1980 and 1984 who were 12 to 17 years old when first interviewed in 1997. Respondents have been interviewed on an annual basis since 1997; round 10 was completed in early summer, and data collection for round 11 began in October 2007. In the 9th round (the most recent for which data are available and completed in summer of 2006), 84% of sample members were interviewed. Interviews are about one hour in length and are conducted face-to-face in around 90% of cases (the rest are by phone.) Sensitive information is collected via a self-administered section of the questionnaire in which the computer is turned over to the respondent so her answers can not be monitored by the interviewer.

In its first 11 rounds, the NLSY97 has followed youths from their early to mid-teens to their early to mid twenties. As such, the focus has been on their development into adults – the on-going process of transitioning from school to employment and from the family of birth to other more independent living arrangements. Because the survey focuses on transitions, survey designers chose to structure the questionnaire around a series of event histories. Complete event histories of education, training, employment, marriage and cohabitation, fertility, and participation in government assistance programs are captured in the NLSY97. Event history data support a variety of powerful statistical and econometric analytic techniques, and are also well-suited to capturing data about respondent's lives in a manner similar to how the respondents themselves conceive of their lives.

To collect event history data, respondents were asked at their first interview for their current status, as well as some amount of retrospective history. For example, interviewers ascertained if they were employed and when they first started working at their current jobs. In subsequent interviews, the interviewer starts by rostering all employers for whom the respondent has worked during the reference period. The respondent is asked to update the status of employers for whom she was working at the last interview and then to report all other employers in any order she prefers. The questionnaire collects the employer name and start and stop date for each applicable employer. The interviewer checks the completeness of the employer list multiple times, reading (or showing) the list to the respondent, and asking her to think of any additional employers to report. The employer roster is then sorted and respondents are asked a detailed supplement about each reported employer in reverse chronological order.

This is the standard approach for all event history collection in the NLSY97: spells are captured by the dates they started and stopped, then details of each spell are collected. The focus throughout is on capturing a comprehensive list of transitions and their timings, so that the status of the respondent at any point can be inferred based on the transitions immediately preceding and succeeding that point. In each subsequent interview, even if the respondent misses one or more rounds, she is asked to go back to the status at the last interview and bring the history forward to the present.

I. Event History Data Does Not Necessarily Mean Event History Calendars

The decision to use Event History Calendars (EHC) involves a tradeoff between flexibility and standardization. EHCs allow interviewers the flexibility to gather the requisite information in a manner that is easiest for the respondent. But it sacrifices a degree of standardization in that all respondents may not be asked questions in the same way. Also, EHC

requires some set up that may prove costly for those without a lot to report. Although the NLSY97 does not employ EHC-style interviewing, some of the experiences of the NLS program in collecting event history data provide insights that are relevant to any collection of such data, regardless of the interview format deployed. Moreover, the NLSY97 questionnaire may be a model for adopting many of the survey methodological principles underlying EHC methods without actually deploying EHC techniques and software.

In contrast to the EHC used on the PSID and other similar applications, the NLSY97 interview uses a highly structured interview. Regular prompts, range checks, and clarifications are often built into the questionnaire, where other instruments might routinely leave these to interviewer discretion or as optional instructions within other question text. Each of the survey's key topical areas, whether involving event histories (employment, schooling, marriage and cohabitation, etc.) or not (household composition, income and assets, etc.) is structured similarly.

A major value of event history data is their potential to sequence events in a respondent's life across domains. A key difference between the NLSY97 and EHC interviews is the extent to which the NLSY97 makes use of events across domains as triggers and prompting cues to assist respondents in dating events. In very few cases does the NLSY97 questionnaire actually ask the respondent the relative sequencing of events; rather, it relies on the collection of absolute dates to infer sequences. Hence, relatively minor errors in reporting absolute dates can cause bigger errors in the implied sequencing.

NLSY97 respondents are asked to provide the month and year of a great many events (and the day of the month for employment events) over the course of the interview. Interviewers are trained to probe using only events that have already been reported in the interview, and not events – however monumental – which may have been discussed only tangentially or outside of

the instrument administration. Also, no information from the self-administered sections are brought into the interviewer-administered portions. This policy makes the structure of the NLSY97 interview relevant. As a rough approximation, the sequence of event history topics in the interview follows as such:

- migration
- schooling
- employment
- arrest and incarceration (self-administered)
- marriage and cohabitation
- childbearing and parenting
- pregnancies (self-administered)
- participation in government programs

In particular, note that childbirth and marriage, which may be highly salient for individuals in their early 20s, come after two of the domains with which transitions can co-occur: employment and migration.

Although the interview formats differ substantially, the NLSY97 questionnaire shares many features with EHC techniques:

- flexible collection of spells as respondents recall them, rather than in a prescribed (e.g., sequential) order
- use of sequential probing – one event in a domain triggers reporting of others
- enforcement of date consistency in many areas. For example, periods not working in the NLSY97 are calculated based on reported employment spells and confirmed by the respondent, rather than collected directly from the respondent. In most domains, the end of a spell cannot precede its start.
- bounded interviewing, anchored by priming respondents with their status at the prior interview.
- generally, start and stop dates of spells are collected, rather than interim month-by-month status.
- the questionnaire is programmed and interviewers are trained to prompt and probe for dates and additional spells by referencing seasons, holidays, and other universal

markers of time, as well as by referencing events that have already been reported in the interview.

In the first several rounds of the NLSY97, the resemblance between the NLSY97 interview and a typical EHC interview was greater. At that time, the NLSY97 deployed a hard-copy calendar which closely resembled the EHC screen and was a successor to the work history calendar that was used for many years on the NLSY79. Broad rows represented each domain, while columns indicated months during the reference period. Interviewers recorded personal and general milestones on the calendar, represented durations of spells using horizontal lines, and used the calendar to help respondents identify additional spells and use context to generate more accurate dates.

An effort to increase and intensify use of the hard-copy calendars revealed that most experienced interviewers were treating the calendar as a paperwork requirement rather than as a respondent aid. Interviewers felt strongly that the calendar assisted very few respondents in recall, since the one year reference period was relatively short, and the number of spells to be reported was generally not large. After maintaining the calendar for two additional rounds only for respondents who had missed an intervening interview, the event history calendar was abandoned altogether in Round 8. Since then, interviewers have had laminated calendars that simply show major holidays and help respondents identify dates within the month (since the NLSY97 employment section collects day as well as month and year information). If they choose, interviewers may record key personal milestones on the calendar to assist respondent recall, but its use in this way varies widely.

The NLSY97 design staff has discouraged up-front collection of major life events outside of the context of the questionnaire. Several of the survey concepts may differ in definition from the definitions of which respondents would naturally think. For example, the NLSY97 asks

about marriage-like relationships (cohabitation) of at least one month where the partners established a single household. The survey also inquires about new children in the respondent's life, including not only biological and adoptive children, but also co-resident step-children. One concern has been that up-front collection of events without reading the project's explicit definitions would result in respondents committing to a series of reports that do not match our definitions, then being reluctant to revise their reports later in the interview when introduced to those definitions.

The benefits from using EHC interviewing techniques may be lower in the NLSY97 than in the PSID and other surveys. Because the vast majority (almost 90%) of NLSY97 interviews continue to be conducted in person, it may be that non-verbal and other visual cues permit a wider range of interviewer assistance to respondents in producing accurate dates.

Moreover, in the NLSY97 interview are conducted annually and few transition occur between interviews. Table 1 shows the number of transitions reported by respondents who completed the Round 9 NLSY97 interview. For respondents who were returning after missed rounds, the table includes only the period since July, 2004.¹ About 95% of Round 9 respondents had a Round 9 reference period beginning after July, 2004. The following transitions are considered: (1) start or end of work with an employer (including gaps within jobs), (2) change of schools or enrollment/disenrollment in school, (3) move from one city or state to another, (4) start or end of a marriage or cohabitation, and (5) birth/death/adoption of a child. The number of domains indicates areas of life in which at least one transition took place.

Table 1: Percent of R9 respondents reporting a given number of transitions within a given number of domains

Number of transitions	0 domains	1 domain	2 domains	3 domains	4 domains	Total
-----------------------	-----------	----------	-----------	-----------	-----------	-------

¹ Round 9 interviews took place between October, 2005 and July, 2006.

0	19.0%					19.0%
1		20.3%				20.3%
2		9.1%	9.2%			18.3%
3		4.7%	8.0%	2.0%		14.7%
4		2.1%	5.6%	2.4%	0.3%	10.4%
5		0.7%	3.2%	1.8%	0.3%	6.0%
6		0.5%	1.9%	2.0%	0.5%	4.9%
7 or more		0.6%	2.2%	2.9%	0.6%	6.4%
Any	19.0%	38.0%	30.1%	11.1%	1.7%	
<hr/>						
All R9 Respondents, R9 reference period falling between 06/2004 and 07/2006						
<hr/>						

Perhaps the most notable result in Table 1 is that the modal number of transitions reported by Round 9 respondents is one. This may be surprising, since the NLSY97 youths were age 21 to 25 at the time of the Round 9 interview, a time of life that Americans tend to think of as filled with change. Indeed, 40% of our respondents had either one change or no change during the year of life discussed in the R9 interview. Another 18 percent had only 2 transitions. The maximum number of transitions is 15, although this is driven by the number of job transitions, which ranges up to 13.² The fact that most respondents have very few transitions to report suggests that for a large portion of our sample, sequencing transitions correctly is a relatively simple task.

Table 2 shows the most common combinations of domains in which transitions occurred, as well as the percentage of the total sample reporting this combination in the Round 9 interview. Again, a domain may be represented by more than one transition.

Approximately 30 percent of the sample is represented in these most common combinations, and employment is clearly the dominant domain in terms of co-occurrence. We note that the restriction on sequential probing only to previously administered sections of the NLSY97 interview does not seem to interfere with the most relevant types of sequential probing;

² The numbers in Table 1 are virtually unchanged if we limit the sample to those respondents who participated in both Rounds 8 and 9. However, the limited sample leaves open the question of whether the number of transitions differs significantly for those who miss an interview.

schooling and migration are both asked prior to employment, and these three sections together account for two-thirds of the combinations listed below. The combinations may quite likely change as the sample ages and enters more fully into childbearing.

Table 2: Common Combinations of Transition Domains

Most common combinations of domains in which transitions occurred	Percentage of total sample reporting this combination
Employment and Schooling	11.2%
Employment and Migration	5.4%
Employment, Schooling and Migration	4.2%
Employment and Marriage/Cohabitation	3.3%
Employment, Marriage/Cohabitation and Migration	2.4%
Schooling and Migration	1.8%
Employment, Schooling and Marriage/Cohabitation	1.4%
Marriage/Cohabitation and Migration	1.3%
All R9 Respondents, R9 reference period falling between 06/2004 and 07/2006	

Both Tables 1 and 2 indicate that it is within-domain that the majority of sequencing issues will be more relevant, and where parallel probing and other cross-domain tools may be less important. Even across domains, the incidence of 3 or more domains exhibiting transitions in a given year is quite rare, so the advantages of extensive cross-domain probing and confirmation within an interview are likely limited in the NLSY97 with annual surveys.

The NLSY97 collects event history data without an EHC. There are obviously both costs and benefits to this decision. Survey designers should consider the length of the recall period, the salience of the events to be recorded, and the frequency of transitions when deciding how structured or flexible their questionnaire will be .

II. Limiting the Collection of Some Details May Be Advantageous

Several factors recommend against capturing every last instance of a phenomenon in a respondent's life. One of these is the time it takes to collect the information, contributing to respondent burden, interviewer labor costs, and often other processing costs. A second is the potential added difficulty for researchers in sifting through all of the additional data, much of which may be of relatively little analytic utility. Thirdly, from a data quality perspective, the possibility of recall error increases when we collect information that may not be substantively important and may be of low salience to respondents. Although it may be interesting to capture information on extremely short jobs or cohabitations, how much information on those short spells is necessary or meaningful? And how well are respondents able to report the minute details of a job they spent one week working over a year ago? Indeed, some of the changes that we have made to the NLSY97 questionnaire have been spurred by respondents complaining that we are asking excessive detail about irrelevant details of their lives, and about which they are unable to recall much.

The NLSY97 has taken significant steps toward collecting comprehensive data in certain domains. Even so, the survey does apply some filters to determine what events respondents are asked to report. Table 3 lists a variety of entities and spells captured in the NLSY97 questionnaire, describing for each what the criteria are for inclusion.

Domain	Eligibility Criteria
Jobs	All jobs, fewer questions asked of jobs lasting < 13 weeks. No limit on number.
Gaps within jobs	All gaps lasting one week or longer occurring entirely within a job spell (i.e., not between jobs).
Marriages	All. No limit on number.
Cohabitations	All cohabitations lasting at least one month or longer, and where a single joint household was established. No limit on

	number.
Migration	Until 2005, only opposite sex cohabitations were collected. From one city, county or state to another; only city, county and state are collected. No limit on number.
Schooling – K-12	All ‘regular’ schools, i.e., leading to a degree. (Summer schools not collected, joint enrollment in community college may not be captured.) No limit on number.
Schooling – Post-secondary	Only enrollment in degree programs, and only terms that were completed. No limit on number.
Children	All biological or adopted children, including deceased and adopted out. No limit on number.
Incarcerations	1997 to 2003: All spells where a respondent was sentenced to incarceration in a jail, prison, detention facility, reform or training school ‘even for one day.’ (Does not include overnight detention <i>prior</i> to arrest and booking.)
	Since 2003, we capture up to 3 arrests in a given interview, following them through sentencing and incarceration (if applicable). For respondents with more than 3 arrests, we capture the first incarceration and current incarceration in the reference period.
Program Participation	All spells by program. All cessation of spells lasting at least one month.

The primary advantage of comprehensive data capture is, of course, comprehensiveness. In addition, it may be that asking respondents to report all spells (or incidents or people) leads to better data quality, since respondents are not asked to apply filters or make calculations that could lead to inadvertent omission of more important spells. Researchers can ignore incidents that are not of analytic interest, but respondents may make more errors in determining which incidents qualify for reporting. Although there is some concern that eliciting many spells increases respondent burden in a way that can dampen future respondent cooperation, we have not found any evidence of such a response in the NLSY97. For example, we considered the 406 respondents who completed the Round 9 interview, but had not completed the Round 8 interview. These individuals completed the Round 10 interview at a rate of about 75%. In logistic regressions predicting Round 10 interview completion, we found no contribution of the

administration time of the Round 9 interview or the number of questions answered in the Round 9 interview.

Table 4 shows the number of spells reported in Round 9 for selected domains. These numbers show only spells reported by respondents who had also completed the Round 8 interview, and so does not reflect the case of large numbers of spells due to long reference periods. As the table indicates, the number of respondents reporting large numbers of incidents in any of these domains is quite low, sufficiently so that the commitment to comprehensive data capture is functionally not very different from collecting a fixed number of spells, say up to 5 jobs or 3 residential moves.

Table 4: Numbers of spells reported by all respondents and low income respondents between NLSY97 Rounds 8 and 9

Type of Spell	% reporting number of spells – all R8 and R9 completers (n=6932)	% reporting number of spells – low income R8 and R9 completers (n=2431)*
Jobs	1 job 6087 (87.8%)	1 job 2049 (84.3%)
	2 jobs 2910 (42.0%)	2 jobs 1006 (41.4%)
	3 jobs 902 (13.0%)	3 jobs 324 (13.3%)
	4 jobs 255 (3.7%)	4 jobs 93 (3.8%)
	5 jobs 69 (1.0%)	5 jobs 28 (1.1%)
	6 jobs 22 (0.3%)	6 jobs 6 (0.3%)
	7 jobs 6 (0.1%)	7 jobs 2 (0.1%)
	8 jobs 2 (0.0%)	8 jobs 0 (0.0%)
	9 jobs 1 (0.0%)	9 jobs 0 (0.0%)
Cohabitations	1 partner 2801 (40.4%)	1 partner 889 (36.6%)
	2 partners 110 (1.6%)	2 partners 41 (1.7%)
	3 partners 9 (0.1%)	3 partners 3 (0.1%)
	4 partners 2 (0.0%)	4 partners 1 (0.0%)
Migration	1 move 1638 (23.6%)	1 move 629 (25.9%)
	2 moves 316 (4.6%)	2 moves 120 (4.9%)
	3 moves 54 (0.8%)	3 moves 25 (1.0%)
	4 moves 7 (0.1%)	4 moves 2 (0.1%)
	5 moves 1 (0.0%)	5 moves 0 (0.0%)

* Low income defined as 200% or less of the poverty level in calendar year 2004. Note that this variable is missing for approximately 18% of respondents. Are we going to delete the 2nd column, given that we haven't talked about low income populations elsewhere?

A hybrid approach may be the one that has evolved in the NLSY97 section addressing college experience. In this section, completed post-secondary terms that are part of enrollment in a degree program are collected. The sequence is as follows:

1. Identify schools in which R has been enrolled since date of last interview.
2. For each such school, ask what degree R pursued there. If no degree program, no further questions.
3. For each school with degree program enrollment, collect the specific terms of enrollment (e.g., Winter 2007) since the last interview, asking R to omit terms currently in progress. Collect school-level information about the R's degree program and progress to date.
4. For each reported term, ask whether the term was completed. If not completed, collect limited information about reasons for non-completion.
5. For each completed reported term in a degree program, collect detailed information about courses, credits, financing, etc.

By taking this gradual filtering approach, the questionnaire restricts the lengthiest, most detailed data capture for the college terms of greatest analytic interest but guides respondents through the reporting process to minimize the risk of inadvertent omission of terms. The instrument also retains some information about terms and schools not asked about to give some sense of the magnitude of these phenomena (e.g., % of terms begun but not completed, or % of youths enrolling in post-secondary institutions outside of degree programs) for some analyses. CP: Check with Rupa, she may have numbers for you on the number of college loops where this comes into play.

Attention must be placed on potential outliers; large surveys always have some. Asking too much detail for short spells may tax the respondent's patience and her ability to recall. Survey designers should think about how much detail will be used by analysts and make sure not to ask more questions than are necessary. Pursuing data that are not analytically tractable may result in avoidable attrition.

III. Dependent Interviewing Can Help Respondents and Analysts

The NLSY97 collects spell data for a number of domains: employment and unemployment spells, spells of program participation, and spells of schooling for instance. Well-known data quality problems arise with spells, particularly those that start in one year and end in the next. Because of the potential problems, the NLSY97, like many other longitudinal surveys, uses dependent interviewing in many sections; that is, information given at the time of the last interview is pre-loaded and fed back to the respondent at the time of the interview.

The practice of dependent interviewing improves data quality in several ways. First, preloading information reduces seam issues by reminding the respondent of ongoing spells at the time of the last interview. In addition, providing the respondent with the opportunity to deny preloaded data helps to avoid problems that might arise from incorrect preloaded data.

In past rounds, the NLSY97 has used two approaches to feed back information to respondents. The first approach consists of reminding respondents of the information they provided at the time of the last interview, and asking them to verify this information (proactive dependent interviewing). For instance, the interviewer asks:

“My computer shows that you were receiving unemployment compensation on [date of last interview]. Is this correct?”

In this case, the respondent is asked to confirm the information given at the date of the last interview.

In the second method, the interviewer feeds back to the respondent the information from the previous interview, but does not ask the respondent to verify this information. The respondent is able to correct the information if he chooses to do so, but the respondent must take the initiative (reactive dependent interviewing).

Beginning with the Round 11 questionnaire, the NLSY97 has moved away from using primarily proactive dependent interviewing toward using the second method of dependent interviewing where respondents are not asked to verify preloaded information, but do have the option of correcting this information. The change occurs in the sections on program participation, schooling, and assets. Survey staff hopes that this change will result in a reduction of respondent burden—both in terms of the time of the interview and in terms of respondents’ perception of being asked and repetitive questions—and reduction in respondents providing contradictory information, while, still permitting the respondent to correct a mistake before continuing the interview.

Second, NLS management takes the view that information collected closest to the time of an event is most likely to be correct. If a respondent claims to be working at a particular job, receiving benefits from a program, or attending a school at the time of an interview, and then denies this event at the time of the next interview, it seems more likely that the information given at the time of the initial interview—when the event was ongoing—is correct. Asking the respondent to verify an event at a later date increases the likelihood that the respondent will deny this event and is more likely to induce a higher rate of error.

Last, when asked to verify previously reported information the respondent may provide inconsistent answers because of confusion over what he is being asked, not because of disagreement with his own previous answer as discussed above. For example, in the NLSY97 Round 7, respondents whose history shows that they have never received unemployment insurance benefits as of the date of the last interview are asked:

“My computer shows that you never received unemployment compensation as of [date of last interview]. Is this correct?”

Of the 6639 respondents asked this question, 178 (2.7%) deny that the computer preload is correct. The 178 respondents who deny are then asked:

“Have you ever received any unemployment compensation payments?”

Of these 178 respondents, 158 (88%) say that they have not ever received any unemployment compensation payments. In these cases the verification increased respondent burden, but did not improve data quality. The response pattern is similar for questions about food stamp receipt, where 194 out of 6510 respondents (3%) deny the pre-load that shows they had never received food stamps. Out of these 194 respondents, 159 (81%) then assert that they have never received food stamps.

In all rounds of the survey, the respondents are reminded of their employment at the last interview, but are not asked to verify that information. The data on employment provide some evidence that this method may reduce the number of times respondents contradict themselves because of difficulty recalling events from the past. Respondents are reminded of all jobs they reported holding at the date of the last interview. They can deny having worked for an employer, although they are not prompted to verify this information. Later in the employment section, they are asked to report the stop date of jobs that were ongoing as of the date of the last interview. Some respondents report a stop date that is before the date of the last interview, indicating that the job was in fact not ongoing. The data show that respondents are much more likely to provide a stop date that occurs before the date of the last interview than to deny holding that job at the date of the last interview.

Year (Round)	# Denying holding Job at date of last interview	# Providing Stop Date for Job before Date of Last Interview
2000 (4)	32	197

2001 (5)	30	139
2002 (6)	22	133
2003 (7)	19	173
2004 (8)	21	208
2005 (9)	15	156

NLS management believes that it is likely that the stop dates provided are a result of recall error.

The reported stop dates are provided to users, but are not used to update the event history data.

For the most part, on topics where the NLSY97 collects event history data, the section begins by confirming the respondent's situation as of the date of the last interview, and then follows by rostering and collecting start and stop dates for events. In at least two instances, point-in-time information is collected before launching into the event history. In the marriage/cohabitation section, respondents are asked to report their current legal marital status prior to confirming marital and cohabitation status at the last interview. In this instance, NLS management believes that asking about current status is helpful for the respondent and for data processing. Asking about current status may help the respondent switch to a new topic before performing the more difficult task of confirming what he was doing at the date of the last interview.

Second, the schooling section starts by asking summary items, such as, highest grade completed, highest grade attended, and expected graduation date prior to collecting the history on schooling. In the majority of the cases, the information that the respondent provides when asked event history questions about schooling (schooling loops) indicates the traditional progression through school. For such cases the NLS program relies exclusively on information provided in the schooling event history to create variables describing education attainment. In more complicated cases where the respondent's reports do not follow the traditional progression, the NLS program uses both the summary variables in addition to the event history reports to

create variables describing educational attainment. In constructing the schooling arrays (a set of variable that provide information on whether and what school the respondent is attending in each month), the created variables on educational attainment from the last and current interviews serve as the start and end values. Once the values for the endpoints have been established the reports from the schooling loops are used to fill in the values in the monthly schooling arrays.

Without much firm data, we believe that reactive dependent interviewing is an important tool in the collection of event history data. It provides a starting point for the collection without giving respondents too much liberty to change answers given in previous interviews. Anchoring the respondent by asking a current status question is also effective and helps adjudicate difficult cases during analysis.

IV. Event History Data Help Maintain Effective Response Rates

One of the key issues for longitudinal data sets is maintaining response rates. Longitudinal data become more valuable as more rounds of the survey are collected, permitting analysts to see the sample members' decisions and outcomes over a longer horizon, but they attain their value only if the sample remains representative.

In the first few rounds of data collection, the rates of wave non-response ranged from 7 to 10 percent for the entire sample. In Round 9, over 18 percent of the sample did not participate, though in Round 10 the non-response rate fell to 16 percent. The patterns of wave non-response are comparable for the whole sample, the oversamples, and the cross-sectional sample.

However, because of the collection of event history data from the date of the last interview, the more relevant statistic for assessing how attrition impacts the NLSY97 is the percentage of sample members who never return to the survey after leaving. Table 6 shows that

most of the NLSY97 sample members who miss an interview return to the survey. Of those who did not participate in the Round 2 interview, less than 25 percent have not returned to the NLSY97 by Round 9. Those who did return were asked about their lives since the last time they were interviewed so that over 98 percent of respondents have now reported on their lives between round 1 and round 2. Because the NLSY97 is able to bring sample members back in each round, their event histories are filled in though the quality of their reports may, depending on the saliency of the event, deteriorate as time passes. This keeps the effective response rate high. Researchers can use the data from non-respondents in a particular round as long as they have returned in subsequent rounds.

Table 6: Response Rate and Rates of “Never Returning” to the NLSY97

Year (Round)	Response rate	Rate of never return among those who 1 st attrited in a given round	Effective response rate
1998 (2)	0.933	0.239	0.984
1999 (3)	0.914	0.288	0.975
2000 (4)	0.899	0.348	0.965
2001 (5)	0.877	0.383	0.953
2002 (6)	0.879	0.508	0.939
2003 (7)	0.863	0.605	0.917
2004 (8)	0.835	0.726	0.880

By estimating the probability of returning to the survey after having missed an interview for the first time, Aughinbaugh and Gardecki (2007) check how those who return to the NLSY97 affect the representativeness of the NLSY97. Their results show that at the last interview having recently had a child born has a significant and positive impact on the probability of returning to the survey. This is good news as these same sample members are more likely to leave the NLSY97 in the round following the birth of the child, and these results indicate that new parents are more likely to return to the survey after initially missing an interview. In some of the samples considered (whole sample, oversample, females, and whites), being employed at the last

interview decreases the chance that the individual returns to the survey. Other life events such as marriage, cohabitation, and being enrolled in school at the date of the last interview are unrelated to the probability of returning to the NLSY97 after initially missing an interview. In addition, the estimates of the probability of return show that the level of the incentive offered increases the likelihood of returning to the NLSY97 by 20 to 30 percentage points for every increase of \$10.³

When the patterns of attrition are examined separately for sample members who either had families with a poverty ratio of less than or equal to 3 in 1996 or had a parent who did not graduate from high school, no notable differences emerge for the disadvantaged subgroup. In addition, when an interaction term between the incentive amount and a dummy variable indicating that one of the criteria from above holds is included, the results provide no evidence that incentives affect participation in the survey—measured by attrition or return—differently for the sample members from less advantaged backgrounds.

With approval from OMB, the level of the incentive offered for participation is a tool at the disposal of the NLS management. Over Rounds 1 through 10, incentive payments range from \$10 to \$50. The incentive level has twice been raised for all respondents, and three incentive experiments have been conducted over the first 10 waves of the NLSY97. In the first three rounds, respondents were paid \$10 for participating in the NLSY97. Over Rounds 4 and 5, the incentive amount was raised to \$20 which remained the payment through Round 8. As an experiment, for half of the sample the incentive was raised to \$15 in Round 4 and then to \$20 in Round 5, whereas for the other half the full increase to \$20 took place in Round 4. A second incentive experiment was conducted in Rounds 7 and 8 in which half of the respondents who had missed the Round 6 interview were offered an additional \$5 per round missed since the last

³ Because incentives increase with round, it may be difficult to interpret the incentive effect independently from the effects on the round indicators. Later rounds are negatively associated with returning to the NLSY97.

completed interview for up to 3 missed interviews in order to compensate for the longer interview. In Round 10, the incentive amount was raised \$30 for all respondents, and a third experiment was launched. In this third experiment, sample members who had not completed the Round 10 interview by December 23rd 2006 were divided into three groups: (1) a control group for which the incentive remain \$30, (2) a cash incentive group which was offered \$50 to complete the interview, and (3) an in-kind incentive group which was offered \$30 cash plus an in-kind gift tailored to the respondent worth an average of \$20. In Round 11, the same incentive offered in Round 10 will again be offered, with the exception that the experimental groups will be expanded to include any siblings who completed the interview prior to the start of the experiment in Round 10.

The variation in incentives in the NLSY97 provides ample opportunity to study how incentives are related to participation. NLS management has examined the results from the experiments, but concedes that the effects of incentives should be studied more, particularly with regard to the effects on respondent behavior over multiple rounds. The effects of incentives may be temporary, constant, delayed, or cumulative. Based on past analysis, we believe that incentives moderately increase response rates and have a greater impact on respondents who have not participated in the past round relative to those who have participated.

Datta, Horrigan, and Walker (2001) study the effects of the Round 4 experiment on Round 4 survey participation. They find that the sample members who were offered \$20 were more likely to participate than those offered \$10 or \$15. However, receiving the \$20 incentive did not differentially impact the most difficult cases, defined based on nonparticipation in Round 2 or 3 or a high number of personal contacts in either of those rounds. The final report on Round 5 of the NLSY97 examines how the second step of the experiment affects participation.

Response rates in Round 5 vary little based on Round 4 payment levels, and those receiving the largest increase in Round 5 have a slightly lower participation rate than do those receiving smaller increases. In sum, the incentive experiment conducted in Rounds 4 and 5 seems to imply that increases in the incentive amount have small increases on survey participation. Both the Round 5 Final Report and Datta, Horrigan, and Walker (2001) examine the effects of the increase in incentive on data quality and find no relationship.

The NLS program examined participation rates in Rounds 7 and 8 for the control versus experimental group where those in the experimental group received an additional \$5 for previous round missed for up to 3 rounds. Participation rates are higher for the experimental group versus the control group. In Round 7 the response rates are 32 percent versus 27 percent, and in Round 8, 36 percent versus 30 percent. In addition, the effect of receiving the experimental fee in Round 7 seems to increase participation in Round 8, even though for such respondents the level of the incentive in Round 8 is below that received in Round 7. Among those who participated in Round 7, 61 percent of the control and 67 percent of the experimental group took part in Round 8. Among those who did not participate in Round 7, 19 percent of the control group and 21 percent of the experimental group take part in Round 8. These results imply that higher incentives help bring attritors back to the survey and upon return, the returnees who were paid more in the past round are more apt to continue to participate in the following round.

The response rates from the Round 10 incentive experiment also imply that higher incentives help bring attritors back to the NLSY97. Participation rates among those who took part in the Round 9 survey are 87 percent for both the cash and in-kind experimental groups and 82 percent for the control group. The differences are larger among the sample members who did

not participate in Round 9 with 37 and 35 percent of the cash and in-kind experimental groups, and 29 percent of the control group participating in the Round 10 interview.

V. Designing Data Files Should Also Be a Priority

Regardless of the structure of the interview, once data are collected in event history format, the question becomes how best to create data files so that researchers can easily and accurately use the data for analysis. The standard method is to present the answers to all questions in the order in which they were asked. In a traditional paper and pencil interview, this has obvious appeal: it is relatively simple to follow along the path the interviewer used to navigate the questionnaire. Computer assisted personal interviews with their complicated skip patterns call into question whether this is still the best method. Event history data adds another layer of complexity, practically demanding that the standard data files be supplemented with other information that eases the researchers' task. Thus, just as a great deal of thought and resources are allocated to the design of the questionnaire, designing the data files should be a priority with survey administrators.

There are several reasons that the standard data files may not be the best delivery mechanism for event history data. First, the number of loops asked varies from respondent to respondent depending on how many changes in status she has experienced. These loops can be fairly complicated, especially if a large amount of information is collected about each spell. Second, the boundary between spells is not always distinct either because they may not have well-defined beginnings and endings (for example, cohabitations) or because they may overlap (jobs). Third, connecting spells across rounds may be difficult because the information needed to differentiate spells may not be releasable. For example, the interviewer may have asked the

respondent about a specific job, “When you were working at Microsoft...,” but when the data are released the identifiers are stripped off. This can make it extremely difficult to connect this round’s job with last round’s job. Finally, sometimes design issues will lead to very complicated patterns that involve changes of status within the spell. The prime example here is capturing event histories of marriage and cohabitation at the same time. It would seem redundant to ask questions on marriage and questions on cohabitation in two separate sections, but following changes in both simultaneously can be very complicated for certain respondents.

The researchers’ needs for event history data may be quite varied. If a researcher wants to estimate hazard models, she may need to draw upon the full richness of event history data. This may include following a respondent’s status on a weekly or monthly basis across several domains. On the other hand, a researcher may want only to know the respondent’s status at one particular time. While each of these needs can be addressed with the event history data, the best presentation of the data on the data files may be very different depending on their needs.

In the early 1980s, the NLS Program realized the difficulty of turning data collected through event history methods into a usable form. In order to perform quality control and create ancillary variables, personnel at CHRR created a file using the employment data in the NLSY79 that laid out the respondent’s employment history on a weekly basis. Originally created for in-house use, this file, termed the Work History data, was made available to the public in 1985. For each week, starting in January 1978, a variable was created to indicate whether the respondent was working, unemployed, out of the labor force, or in the military. Other weekly variables tracked the total number of hours worked and jobs that were held concurrently. Each job was given a number that could be used to track it across rounds or across periods of layoff. These job numbers can also be used to track other information about the job such as hourly rate of pay,

industry, and occupation including summary variables for that particular job that are often calculated from the Work History data. Finally, summary variables that aggregate information across all jobs (total weeks worked, unemployed, or out of the labor force) are present both by year and since last interview.

From the survey's inception, the NLSY97 design team decided to increase the amount of data collected in event history format. Because of this, researchers know what is happening in the lives of respondents across several domains. The data files present these data using three types of special variables. The "event history arrays" give the status of the respondent in terms of employment, schooling, marriage and cohabitation, and government program participation on a weekly or monthly basis. These can be linked to "spell data" that describe the situation during a given spell in greater detail (for example, job or spouse characteristics). The NLSY97 also relies heavily on "created variables" that summarize the information in the event history record. These often tell the status of the respondent as of the date of the interview or the number of status changes between interviews or over a longer period of time. All three types of these variables could be calculated from the event history data, but many would require utilizing hundreds of variables in order to make the calculations. Having the survey management team perform these calculations ensures accuracy and consistency across researchers.

To make it easier to track and determine the amount of time between events, a monthly numbering system was introduced for all monthly event history arrays. January, 1980 was designated as month number 1 and subsequent months were numbered sequentially from this month. Consider a woman who was born in month 15 (March 1981), first married in month 294 (June 2004), and had her first child in month 308 (August 2005). The monthly numbering system makes it easy to calculate age at first marriage (279 months or 23 years, 3 months), age at

first birth (293 months or 24 years, 5 months), and time between first marriage and first birth (14 months). The one exception to this numbering system is employment data, which are reported weekly. Unfortunately, matching weeks to months is not totally straightforward (some rule must be devised to account for weeks that straddle two months), but the greater frequency of job changes makes weekly data necessary in the employment area.

As one might imagine, creating variables to report monthly or weekly status as well as summary information about spells and status over time has led to a proliferation of variables in the data files. Over 20,000 new variables are now added to the NLSY97 every round. This obviously creates burden on data users who must wade through all these data. Currently, NLS management is in the process of classifying all variables in its surveys into three classes according to how likely researchers are to use the variable. Created variables and event history variables are generally designated primary variables indicating that these are the variables that best summarize and describe the event history data. These are especially useful to researchers who are not focusing on the event history variables for their analysis but would like to use this information as control variables. Researchers who are focusing on event history variables may want to delve into the secondary variables that include the raw sequence of questions that are the basis of the event history data. Computer checks and little used variables make up the third level of variables. It is hoped that this classification of variables will make it easier for analysts to find the variables they need more quickly.

The experience of the NLSY97 has been that careful thought must be put into how data files are created. As CAPI instruments become increasingly complex, it is not enough to simply record the answers to each question asked. The NLSY97 approach has been to create a whole series of data – event history arrays, spell data, and summary created variables – while

maintaining the basic raw data for the most advanced users. Other approaches may work even better. The key is to recognize the difficulties event history questionnaires present to the users and to plan for data files that will alleviate these problems as much as possible.

VI. Conclusion

The success of any longitudinal data sets rests on the choice of sample members to participate and on the choice of researchers to study the data. In the design and implementation of the NLSY97, the NLS program attempts to balance the interests of these two groups, which in some instances pull in opposite directions.

The NLS interviews are more structured than the EHC method used for the PSID and other surveys for three reasons. First, the NLSY97 interviews were designed in 1996, before EHC interviewing was commonly used. For consistency the structure of the interview has remained the same. Second, the format of the NLS interview ensures standardization across all respondents in how the interview is administered. Third, given the short recall period, the salience of the events, and the fact that interviews are primarily conducted face-to-face, it is not clear that using the EHC methodology would provide higher quality answers or reduce respondent burden in the NLSY97. Nevertheless, we feel that much that we have learned from the NLSY97 is applicable to any survey that collects event history data.

One decision that we have made is to not to collect all details on every spell for every event history. This decision reflects our experience that many respondents have difficulty recalling details of events that occurred far in the past and lasted a very short time. Collecting information on these events would increase respondent burden without improving data quality. A second lesson we have learned is that, while dependent interviewing improves data quality, asking respondents to confirm all information from the past does not improve data quality but

can increase respondent burden. As the NLS program considers mechanisms for improving the reporting task or quality of reports from the remaining sample, the program must weigh the costs and experiences of the relatively straightforward majority.

One of the advantages we have found of using the event history format is that it reduces the impact of attrition on the data. Most respondents who miss an interview return in subsequent rounds. This allows us to collect information to complete their event histories for the missed rounds and increases the usefulness our data for researchers.

Once data are collected in event history format, the question becomes how best to create data files so that researchers can easily and accurately use the data for analysis. The standard method of presenting the answers to all questions in the order in which they were asked may not be the best for complex event history data collected with CAPI methods. To make the data more accessible to users, we have created whole series of data – event history arrays, spell data, and summary created variables. We are in the process of re-organizing the way researchers view the data and hope that the new format steers them to our created variables. We also continue to investigate other approaches to make complex data more usable.