# Bridging the Data Gap for Marriage and Family Research: Potential Opportunities Within the NLSY97

December, 2006

**Prepared for:**

Administration for Children and Families
Department of Health and Human Services

**Submitted by:**

Kelleen Kaye
Senior Research Scientist
National Opinion Research Center

NORC
AT THE UNIVERSITY OF CHICAGO
A national organization for research and computing

## OVERVIEW

**Introduction.** While the scope of research on marriage and family formation has expanded greatly since the 1970s, the basic need to understand how families are doing, what challenges they face, and what helps them thrive will continue to be important. Marriage-related studies have evolved from merely tracking trends, to describing pathways into relationships and parenthood, to analyzing influences on child well-being and informing the Healthy Marriage Initiative. As a result, the field today covers an array of related topics including marriage, the wider spectrum of family structures, fatherhood, community resources, social networks, and the role of policy and programs as they relate to family well-being. There is a growing body of research showing how each domain influences family well-being directly and is vital in its own right. However, in combination they influence family well-being in interactive ways that are still not fully understood, for example how some elements mitigate or magnify the influence of others and how their relative importance varies over the life course.

Policy makers and researchers need to better understand how these dimensions of the family context intersect, and what this implies for developing policies and programs to strengthen families. To help achieve this, we need data that track individuals into relationships and parenthood; examine interactions among family members inside and outside the household; describe family resources, stressors and well-being along multiple dimensions and points in time; catalogue program participation; and capture a wide array of related covariates. In addition, we need data that will allow for examination of the roles and implications of these factors among different understudied populations such as low-income families and racial and ethnic minority groups.

This paper discusses the richness of the National Longitudinal Survey of Youth, 1997 Cohort (NLSY97) for studying these issues, and ways in which its utility for advancing research on marriage and the family could be enhanced. The basis for the recommended improvements comes from the discussion of a panel of experts convened by NORC for the Administration for Children and Families. To put these recommendations into clearer perspective, this paper begins with a review of the promises and current limitations of NLSY97 for studying marriage and family issues and ends with a discussion of first steps one could take in pursuing such enhancements.

**The Utility of NLSY97 for Marriage/Family Research.** NLSY97 is one of the most promising data sets to embody the critical characteristics needed for marriage/family research. It details the family circumstances of adolescents who were ages 12 to 16 in 1997. As a longitudinal data set, it has followed the transitions of these adolescents into early adulthood, and will continue to track their activities, relationships and well-being each year as they mature.[1] The data are nationally representative with an original sample of nearly 9,000 observations, and include minority over-samples. There is extensive detail on family circumstances, not only for families headed by the youths' parents but also, later, for families headed by the youths themselves. This detail includes: measures

[1] The current contract calls for an additional five rounds of data, and there are expectations for future extensions of the data as well, similar to the NLSY79 cohort which continues to be funded after 21 rounds.

of relationship status (e.g., married, cohabiting, dating); parental status (e.g., residential, biological, marital, adoptive); data on nonresident parents, visitation and child support; relationship quality and family processes; measures of wellbeing; and detailed covariates in domains including education, work, income and program participation, risky behaviors, health, attitudes/expectations, and social context. Key transitions in many outcomes are also captured in extensive event history modules.

**Opportunities to Learn More.** These data offer a rich portrait of families over time and across two generations, yet their potential to inform marriage and family research is largely untapped due to their complexity and other limitations. With this in mind, the Administration for Children and Families (ACF) asked NORC to convene a panel of experts to explore ways to strengthen the utility of the current NLSY97 for marriage and family researchers. Experts were asked to discuss these limitations and identify opportunities for enhancing its use in the future in relation to what they felt were important unanswered questions in the following areas:

- Pathways to cohabitation and marriage
- Relationship between family formation/marital status and employment
- Marriage outcomes
- Child well-being in different family structures

Within these topical areas, panel members felt the data offered significant potential to tackle what they felt were the critical unanswered questions in the field. The panel members' suggestions for improving the data's utility include several that would enable researchers and policy makers to more easily tap the NLSY97 as a ready source of information on how today's families are doing. A second group of suggested enhancements would form a platform for greater collaboration on marriage and family research across policy areas and organizations, and establish a common body of knowledge that could help to leverage advancements in the field. Finally, the panel suggested enhancements they felt were vital to expanding our understanding of family functioning beyond what is already known, and preparing for the key family research questions on the horizon. These suggestions include the following:

*Transforming the NLSY97 into a Ready Reference on Marriage and Family:*

- Develop a user-friendly research roadmap--i.e., targeted NLSY97 documentation on marriage and family topics, with detailed information on key measures and variables, caveats and limitations, and control totals for populations of interest.
- Provide an on-line table generator that intersects marriage and family variables with other domains, such as work, education, health and program participation.
- Assess the adequacy of sample sizes and response rates for specific populations of interest (subpopulations suggested by the panel include low-income respondents, racial/ethnic groups, those who have been incarcerated or young black men.) Construct a "flat file" that would contain a limited subset of variables on the most relevant topics related to marriage and family in a user-friendly format.
- Provide more created variables that easily identify complex relationships, such as whether the respondent's cohabiting partner is also the biological parent of her child.

- Identify, for key variables, the universe of eligible respondents, so the implications of missing data can be more easily assessed.

*Creating a Common Knowledge Base and Platform for Collaboration/Leveraging:*

- Establish an on-line collaboratory or message board on NLSY research topics related to marriage and family. This could be used to discuss difficult constructs, such as how to best define cohabitation, or to better encourage marriage and family research that reaches across multiple policy domains.
- Provide a repository for computer programs, tabulations or definitions of key outcomes related to marriage and the family within the NLSY97.
- Fund and publish analyses on key marriage-related questions and their policy implications using the NLSY97. This could include awarding dissertation or other small grants.

*Deepening Our Understanding of Families and Planning for the Next Generation:*

- Add survey questions on topics such as relationship skills, domestic violence, power and resource sharing, work/family balance, financial literacy, reasons why unmarried, measures of trust, and participation in or awareness of healthy marriage services.
- Capture spheres of influence outside the household, particularly as they affect children. This could include asking some questions of the focal respondent's partner or the nonresident parents of their children, as well as extended family members.
- Add new respondents, such as immigrants, to make sample more reflective of today's young adults.
- Add new questions that provide a more complete rostering of multiple partner fertility and complex step or social parent interactions.
- Add new questions that measure respondents' interaction with their children, and their children's wellbeing.
- Plan for the development of new cohorts. One proposed cohort representing a new cross section of youth in 2010 is already the topic of serious discussions between BLS and other government agencies, including the Department of Defense.

**Overview conclusion.** The panel of experts convened during the ACF-sponsored workshop was impressed by the depth of content in the NLSY97 as a significant resource for research on marriage and family. Proceedings from the workshop made it clear that the NLSY97 data could greatly inform critical unanswered research and policy questions in the area of marriage strengthening and child and family well-being. For example, the data could suggest answers to questions directly related to the ongoing implementation efforts of healthy marriage interventions such as:

- What are major barriers to forming and maintaining healthy marriage in society?
- What relationship skills and qualities appear most critical for healthy marriage?
- What are the differences among couples (e.g., by income or race/ethnicity in types and qualities of relationships and family structures?

- Which couples are at greatest risk and might benefit most from healthy marriage interventions?
- What sources of challenge and resilience among couples appear most relevant for healthy marriage and child well-being?

The data could answer similarly critical questions that intersect marriage with other important spheres of family functioning and highlight potential implications for policy more broadly. For example:

- What can be learned about how some families thrive when facing adverse conditions? What contributes to movement of disadvantaged families up the economic income scale? What role do public programs and social structures play?
- How can we better understand today's complex family structures and identify ways to help ensure that all parents (biological and step, resident and non-resident) are appropriately engaged in the child's wellbeing?
- How can we better understand the relationships between decisions regarding family formation and child bearing and other critical decisions such as those regarding employment opportunities or living arrangements/housing and the implications for each domain?
- What points over a child's life appear to offer the greatest opportunities for families and society to make a difference in their well-being, and at what points are they most vulnerable?

The next section reviews the characteristics, strengths, and current limitations of the NLSY97 data. This is followed by a summary of the potential enhancements recommended by the expert panel, and a discussion of possible next steps in pursuing those enhancements.

## The NLSY97 Data, Its Strengths and Limitations

**Background of the NLS Program and Structure of the NLSY97 Cohort.** The National Longitudinal Survey of Youth is part of a broader collection of surveys sponsored by the Bureau of Labor Statistics within the Department of Labor, and collectively known as the National Longitudinal Surveys, or NLS. Their primary function is to gather information at multiple points in time on the labor market experiences of groups of men and women. The data are longitudinal, meaning that once a cohort is sampled the selected individuals are then re-interviewed (either annually or biennially) over several years. The NLS includes a total of eight cohorts of men and women. Four of these were selected in the 1960s and those surveys have since ended; the four more recent cohorts continue to be interviewed. Thus, for several panels of men, women and children, the data capture a significant segment of the individuals' lifetime, while also reflecting detailed changes from one year to the next. The table below summarizes these cohorts as described in the official NLS Handbook of 2005, including definition of the cohort by age and birth year, original sample size, and the span of years covered by the interviews.

**Cohorts of the NLS**

| Survey Group | Age cohort | Birth year cohort | Original sample | Initial year / latest year | Number of surveys | Number at last interview | Status |
|---|---|---|---|---|---|---|---|
| Older men | 45-59 | 4/1/06-3/31/21 | 5,020 | 1996 / 1990 | 13 | [1]2,092 | Ended |
| Mature women | 30-44 | 4/1/22-3/31/37 | 5,083 | 1967 / 2003 | 21 | 2,237 | Ended |
| Young men | 14-24 | 4/1/41-3/31/52 | 5,225 | 1966 / 1981 | 12 | 3.398 | Ended |
| Young women | 14-24 | 1943-1953 | 5,159 | 1968 / 2003 | 22 | 2,859 | Ended |
| NLSY79 | 14-21 | 1957-1964 | [2]12,686 | 1979 / 2004 | 21 | [3]7,724 | Continuing |
| NLSY79 children | birth-14 | - | [4]- | 1986 / 2004 | 10 | [3]3,190 | Continuing |
| NLSY79 young adults | [5]15 and older | - | [4]- | 1994 / 2004 | 6 | [3]4,238 | Continuing |
| NLSY97 | 12-16 | 1980-1984 | 8.984 | 1997 / 2004 | 8 | [6]7,756 | Continuing |

[1] Interviews in 1990 also were conducted with 2,206 widows or other family members of deceased respondents.
[2] After dropping the military (in 1985) and economically disadvantaged nonblack/non-Hispanic oversamples (in 1991), the sample contains 9,964 respondents eligible for interview.
[3] The latest sample size available is from the 2002 survey.
[4] The size of the NLSY79 child sample depends on the number of children born to female NLSY79 respondents, attrition over time, and the gradual aging of the children into the young adult sample. The size of the young adult sample depends on the number of children who reach age 15 in each survey year. Information about the number interviewed in each survey is available in chapter 4.
[5] In 1998 only, the young adults eligible for interview were limited to those ages 15 to 20.
[6] The latest sample size available is from round 7.

The most recent cohort, the NLSY97, is the focal point of this paper. Its purpose as stated by BLS in the NLS Handbook is to "document the transition from school to work and from adolescence to adulthood." As such, it is representative at the national level of individuals born between 1980 and 1984, i.e., between the ages of 12 and 16 in 1997. The sample includes 8,984 members, including an over-sample of just over 2,000 additional blacks and Hispanics. The progression from Round 1 to Round 8 reflects a high level of sample retention, just over 86 percent, which is typical for surveys in the NLS program.

It is important to note that not all these respondents represent unique households. The NLSY97 interviewed all adolescents of qualifying age within a household. As a result, just over 5,000 individuals were from single respondent households and just under 4,000 were from multiple sibling households. The inclusion of siblings as respondents represents a unique opportunity to study variability of outcomes within a household. BLS recommends caution when generalizing from studies of siblings within the NLSY97, however, because these households are not representative of all sibling households.

In addition to the annual youth questionnaire, the NLSY97 contains special data supplements in various rounds. These include a detailed interview of a youth's resident parent in round 1, a household income update completed by a residential parent in rounds 1 through 5, the collection of high school transcripts in 2000 and 2004, surveys of secondary schools located within the primary sampling units in 1996 and 2000, and test scores for the *Armed Services Vocational Aptitude Battery (ASVAB)* in 1997. In addition, the data contain screener information for all members of the household in round 1 and roster information for household members and immediate family outside the household in all rounds.

**Topical Content.** Although the study of labor market outcomes is a primary goal of the NLSY97 survey, it also contains a rich set of covariate outcomes for the respondents--pertaining to schooling, training, income, assets and program participation, family formation and family background, expectations, attitudes, behaviors and time use, health, environmental variables and event history variables. For a full description of the array of variables within each of these domains see the *NLSY97 Users' Guide*.[2] Here we briefly highlight the content of particular interest to the marriage and family formation agenda.

A recent ACF review of data sources available to conduct research on marriage and the family concluded that the NLSY data are well-suited for measuring a variety of factors related to family formation, including economic, cultural and demographic influences.[3] Among several valuable data sets reviewed, it is shown to be the only survey that combines rich measures of relationship quality with detailed relationship histories (including cohabitations), extensive employment and human capital measures, and

[2] The *NLSY97 Users' Guide* is available on-line at http://www.bls.gov/nls/97guide/nls97usg.htm.
[3] See "Marriages and Family Formation Data Analysis Project," available on-line at http://www.acf.hhs.gov/programs/opre/strengthen/marr_family/reports/determinants_guide/determinants_guide.pdf

additional well-being outcomes with annual interviewing and a nationally representative sampling frame.

Specific details on NLSY97 content are presented below within the four marriage-related topics presented to the panel for discussion:

*Pathways to cohabitation and marriage:*

- Detailed annual questions on cohabiting and marital partners have been assembled into event history variables that track relationship stability and transitions.
- Demographic variables on the other biological parent of the respondent's child, even if they have never cohabited or married.
- Variables are included that measure expectations of marriage within the coming year.
- Marital history of youths' parents captures intergenerational influences.

*Relationship between family formation/marital status and employment:*

- Detailed employment data on all jobs held by the respondent include weekly employment status, as well as benefit eligibility and average weekly hours and earnings.
- Job descriptions include measures of industry, occupation, employer size, circumstances of hire, level of satisfaction and union status.
- Employment gaps are described in terms of timing, reasons and job search efforts.
- Other covariates possibly related to selection into marriage are included, such as education, training, and aptitude scores (ASVAB).
- Related outcomes, such as child-care use by type of arrangement and welfare receipt, are also covered in detail.
- The event history structure allows an analyst to capture timing of transitions more precisely, to better enable estimation of temporal relationships.

*Marriage outcomes:*

- Extensive measures of relationship quality between youths' parents (or parent figures), both as reported by the youth and as reported by the responding parent, include supportiveness, degree of conflict and conflict resolution. Youth-reported measures are collected over multiple years.
- As youth transitions into young adulthood, there are similar measures of relationship quality between him/her and his/her partner are included as well.
- The data support detailed breakouts by marital status of parental interactions with youth--including measures of family routines, parental engagement, decision making and limit setting—and contain information on contact with nonresident parents.

*Child wellbeing in different family structures:*

- Youth well-being is measured across multiple domains, including educational attainment, employment, obesity, risk-taking behaviors and emotional well-being, measured at multiple points in time. Parents also provided more detailed information on youth health status during round 1.

- The analyst is able to observe variation in youth wellbeing across multiple family structures--including married step, married bio, cohabiting step, cohabiting bio, adoptive and single-parent households.
- Parents' retrospective information allows researchers to measure impact on children of total time spent during childhood in married versus nonmarried households, as opposed to capturing only a point-in-time relationship.
- Future data collection will offer the potential to observe well-being of the youths' children and to observe changes over the life-course.

**Data Access and On-Line Resources.** There are currently three methods for obtaining the NLSY97 data. (1) The user may search and extract variables from the NLSY97 using the on-line extraction tool known as Web Investigator, available through links on the BLS website. This tool contains documentation on each variable, including number of observations and missing values, as well as links to the NLSY97 user guide. (2) One can order a CD that will contain all cohorts of the NLS as zipped files, along with extraction software developed by the Center for Human Resource Research (CHRR) at Ohio State University. (3) One can download a particular cohort, which is accompanied by a copy of the extraction software to install on the user's hard drive. Additional details on data access can be found in Appendix A of this paper. User guides and user handbooks are also available for download from the BLS website.[4]

**Current Limitations of the NLSY97.** In addition to its many recognized strengths, the NLSY97 has several limitations in its utility for marriage and family research. These limitations, which are summarized below, are less a sign of weakness in the data than a function of the extensive capacity of the data, plus the fact that the survey was originally developed to study labor market issues. Some limitations particularly relevant to the data's utility in the topical areas discussed above include the following:

- ***Richness of data.*** The sheer volume of measures within the NLSY97 data makes it a powerful tool for studying marriage-related outcomes, but also lends great difficulty to identifying particular measures of interest. For example, an on-line search of the NLSY97 variable names for the word 'cohabitation' yields nearly 600 variables.
- ***Ambiguity across generations.*** The content on family formation outcomes for multiple generations is a powerful benefit of the NLSY97. But variable names often inadequately distinguish between them (e.g., it can be difficult to tell whether a variable is referring to the youth's father or the father of the youth's child).
- ***Lack of specificity in Users' Guide.*** Given the broad array of outcomes covered in the NLSY97 Users' Guide, it can only touch on some topics such as family formation in general terms. While it describes the types of outcomes covered, there is little detail on specific variables and how to use them.
- ***Skip patterns that are complex to disentangle.*** Because the questionnaire is extensively tailored to respondent circumstances, the same construct may be collected from different respondents across a number of different variables. As a result, numerous variables have relatively small numbers of valid responses. When a user

[4] Links to data and documentation can be found at http://www.bls.gov/nls/home.htm

discovers a variable with a large proportion of missing data, uncovering the implication for one's analysis can be quite time consuming. As there is no consistent information on the universe of eligible respondents for each question, identifying why certain observations are missing requires one to trace back through skip patterns across several preceding variables.

- ***Key measures that are hard to calculate.*** Several measures fundamental to the family formation agenda can be calculated based on the existing data. This requires extensive effort as there are no corresponding created variables. A primary example of this is whether the youth's current cohabiting partner is the father of the youth's child.
- ***Lack of interview questions on other key outcomes.*** Key outcomes not addressed include measures of in-kind contributions by the nonresident parent of Respondent's child, welfare receipt by Respondent's mother beyond 1997, non-cash assistance Respondent received from family if not living at home (other than child care, gifts and allowance), and wellbeing measures for child of Respondent.
- ***Lack of control totals on some key subpopulations.*** As noted above, several key subpopulations are difficult to isolate due to the lack of a single identifying variable. The lack of control totals for these groups further limits researchers' ability to have confidence in their estimates.
- ***Limited information on parenting of young children.*** By definition, the respondents to this survey were already teens when they were first interviewed. Thus, we lack information on how their earlier childhoods were affected by the type of parenting they received. At this point we also lack information on parenting behaviors of the youth themselves because many are just now entering adulthood.
- ***Inability to support sub-national estimates.*** With initiatives at the state and local level to promote healthy marriage, contextual measures of family formation outcomes at the sub-national level would be valuable. However, sample sizes within the NLSY surveys are only sufficient to represent national-level populations. This is also true of most other surveys containing the same level of detail, e.g., the National Survey of Family Growth and the Survey of Income and Program Participation.

## Expert Panel Workshop Discussion

With these limitations in mind, ACF convened a one-day workshop that brought together a wide range of experts to discuss how the utility of the NLSY97 data might be strengthened for research on marriage and the family. Attendees included both individuals with extensive experience using the NLSY97 data and members of the NLSY97 design team, as well as experts within the marriage and family formation arena who are typically less familiar with the NSLY data base. The workshop focused primarily on the four topical areas already noted:

- Pathways to cohabitation and marriage
- Relationship between family formation/marital status and employment
- Marriage outcomes
- Child well-being in different family structures

A panel discussion was held on each topic, focusing first on what important substantive questions remained unanswered. Building on this, workshop members then discussed the data limitations within the NLSY97 that have particular relevance to these questions and the steps that would offer the greatest promise for addressing them. Workshop participants were asked to suggest small, medium and large-scale targets of opportunity for enhancing the utility of the NLSY97 with respect to these topics. Below are the major themes that emerged from the discussions. As the four topical areas listed above are inter-related, many of the discussion themes raised by the group came up during more than one panel; for brevity, they are summarized only once.

**Pathways to Cohabitation and Marriage.** The longitudinal nature of the NLSY97, along with its event history module, provides for a detailed tracking of individuals into and out of cohabiting and marital relationships. The same level of detail on important covariates such as fertility, employment and public assistance receipt allow researchers to study the associations among these outcomes, based on the appropriate sequencing of events and temporal relationships.

While workshop participants agreed that the event history data provided an excellent opportunity to track relationship pathways, they felt the data would be stronger with the addition of direct survey questions on key transitional events. One cited example was to ask a respondent directly if she was already pregnant when she married her partner. Another was to enhance the ability to track pathways by further work in defining cohabitation, e.g., measures that would distinguish between cohabitors who clearly saw themselves on a pathway toward marriage versus those who did not.

In addition to tracking and defining relationship pathways, participants discussed the field's ability to understand factors influencing these pathways. They agreed that while the sequencing of events was a critical piece of this picture, important contextual factors were missing from the data. These include measures of the local environment, normative measures among peer groups, participation in or awareness of marriage education programs, and reasons for not engaging in such activities as marriage or parenthood.

**Relationship between Family Formation/Marital Status and Employment.** The extent to which employment fosters marriage opportunities or marriage fosters employment gains has been extensively studied in earlier cohorts of the NLSY and, as mentioned above, can be studied in detail with the NLSY97 as well. However, workshop participants felt that in addition to looking at the connection between marriage and employment, a third element of this dynamic—that of resource and power sharing—was much less understood and missing from the NLSY97 data. The extent to which employment and earnings patterns influence these sharing dynamics, and the interconnections between these dynamics and marriage, remains a largely unanswered question that could be addressed with the NLSY97 given a few additional measures. Participants also felt it would be valuable to have more measures documenting efforts in or difficulties associated with balancing work and family life, and measures of financial literacy.

**Marriage Outcomes.** During this panel, workshop participants discussed outstanding questions and data issues pertaining to the marital relationship itself, including measures of relationship quality, stability and commitment, as well as parental interactions with children. The NLSY97 contains a rich array of variables measuring the quality of relationship between the youth's parents and interactions between youth and parent (both residential and nonresidential), as well as the quality of relationship between the youth and his or her partner in adulthood. This information includes both objective measures based on particular behaviors and subjective measures such as levels of satisfaction.

In addition to these measures of quality, workshop participants talked extensively about the importance of assessing relationship skills, which currently are not captured in the NLSY97 data. There was a strong shared sense that such measures would be valuable to understanding family processes and informing healthy marriage interventions. The specific suggestions for possible directions to take were quite varied, suggesting that additional discussions and conceptual work would be a necessary first step.

In terms of relationship stability, participants noted that while the duration of relationships can be tracked over time, measures of relationship security at a given point in time are lacking. Participants suggested that adding measures of perceived relationship exit costs, shared goals for the future, and worries over infidelity could help address this gap. Domestic violence was also identified as a critical domain that is completely absent from the NLSY97 data. The data could be further enhanced, participants felt, by gathering all measures--existing and newly proposed--from the perspectives of both partners rather than just one.

**Child Well-Being in Different Family Structures.** The NLSY97 data provide a significant opportunity to study the relationship between family structure and child well-being, owing to the extensive range of youth well-being measures (including outcomes related to health, education and risk-taking behaviors), the variety of structures within which the youth are raised (e.g., single versus two parent, biological versus step parent, married versus unmarried parent), and retrospective variables that document how much time the youth's parents spent in each type of relationship.

As family structure becomes more complex, however, it intersects with child well-being in a number of ways that are not adequately captured in the NLSY97 data. Most notably, workshop participants felt that more information was needed on interactions outside the household, particularly regarding time the youth spent with nonresidential parents, social parents or extended family members. On a related note, panelists stated the importance of having a more complete rostering of parental/child relationships both inside and outside the household, including the relationships between a parent's partner and each child as well as the presence of the partner's children outside the household.

In addition to being able to assess how child well-being is affected by a complex family landscape, studying how it changes over the lifecourse is equally important and complicated. As the longitudinal cohort of the NLSY97 enters young adulthood and, for many, parenthood, it is well positioned to study this dynamic as their children age.

However, measures of their parenting activities and their children's well-being have yet to be developed, and this will be a critical next step.

The NLSY97 also offers the opportunity to study this dynamic over the course of the youth's childhood as it relates to changes in their own parents' relationship status, based on the relationship history information provided by parents in round 1 of the survey. To fully realize this potential of the data, however, the panel noted that additional measures would be needed to capture earlier points in the youth's childhood, both in terms of the youth's well-being and as covariates for the parents that may have jointly influenced family well-being and family formation.

**Cross-Cutting Themes.** In addition to issues related to the four topical areas described above, other themes were addressed that were clearly of broad relevance to nearly any NLSY research on marriage and the family. These include the value of: repackaging the existing data and documentation to make it more transparent and usable to policy researchers; supplementing the sample to reflect new waves of immigration; and the creation of new cohorts, possibly a sample of children born to youth of the NLSY97 cohort or a new cross section of youth in 2010.

## Potential Opportunities within the NLSY97 and Next Steps

A number of potential opportunities for enhancement were recommended during the workshop that could significantly improve the ability of the NLSY97 to inform marriage and family research. These include several that would enable researchers and policy makers to more easily tap the NLSY97 as a ready source of information on how today's families are doing. Other changes or new efforts would form a platform for greater collaboration on marriage and family research across policy areas and organizations, and establish a common body of knowledge that could help leverage advancements in the field. Finally, the panel suggested efforts they felt were vital to support movement beyond our current understanding of family functioning and prepare for the key family research questions on the horizon and the next generation of families.

In the table below and the discussion that follows, we follow up on the panel's recommendations to provide a general sense of the efforts one would need to pursue these types objectives. Our description of potential next steps is suggestive, meant to serve as a platform for those with direct responsibility as well as members of the NLSY-user community to consider what range of efforts may be well-suited to their mission and means, rather than as a definitive proposal to complete such tasks.

Many of the suggested efforts would need to be explored in partnership with BLS, the sponsoring agency, particularly those involving content or delivery of the data. Others, such as targeted documentation, could go forward independently, although BLS would still be a valuable partner. These caveats notwithstanding, the illustrations below suggest the data activities agencies or organizations interested in family research and policy might wish to consider, and the steps that could take them in that direction. We discuss each of the major categories shown in the table, and the items within them, in turn.

**Potential Opportunities to Increase the Utility of the NLSY97 Data**

| Opportunity | | Approx. Timing | Type of Effort | Possible Approaches |
|---|---|---|---|---|
| **Transform NLSY97 into Ready Reference on Marriage and Family** | | | | |
| | Create targeted users' guides on marriage and family topics. | Under 12 months. Could be initiated at any time. | Small effort for each topic | Pursue through BLS and its contractors, or convene group of data and subject experts. |
| | Create on-line table generator of family variables and covariates. | Under 12 months. Could be initiated at any time. | Medium | Attach to WebInvestigator, through BLS contract, or create alternative web site. |
| | Construct flat file of key marriage and family outcomes w/covariates. | 12 to 24 months. | Medium | Pursue through BLS and its contractors, or convene group of data and subject experts. |
| | Assess sample adequacy for key populations. | Under 12 months. Could be initiated at any time. | Small | Could be done by any NLSY97 data expert. Suggest getting input from BLS or its contractors. |
| | Provide additional created variables to identify key relationships. | Under 12 months. Could be initiated at any time. Need to start 6 months prior to next data release. | Small effort for each set of variables. | Users can readily create and share variables on their own. Or variables can be developed and added to official data by BLS and its contractors. |
| | Define universe of eligible respondents by variable. | Computer Assisted Phone Interview Feasibility study: under 12 months.<br><br>Implementation: up to 24 months. | CAPI Feasibility Study: Small<br><br>Implementation: Large | Expert data users could do manually for select variables. Theoretically could be done globally as CAPI software enhancement through BLS and its contractors. |
| **Create Common Knowledge Base and Platform for Collaboration/Leveraging** | | | | |
| | Create on-line collaboratory or message board. | Under 12 months to start-up. Could be initiated at any time. | Small effort to start up | Link to NLSY web through BLS and its contractors, or create alternative forums. |
| | Create on-line repository for code, tabulations, etc. | Under 12 months to start-up. Could be initiated at any time. | Small effort to start up | Work through BLS and its contractors to add to NLSY website, or create new site. |
| | Perform analyses of key topics using available data. | Ongoing, but key questions remain unanswered. | Small | Fund work through special journal issues, conferences, dissertation grants, etc. |
| **Deepen Understanding of Families and Plan for the Next Generation of Research** | | | | |
| | Add key survey questions. | Design work must start 2 years prior to next data wave. | Variable | Proposals can be sent to NLSY97 PI. Must be done through BLS contract. |
| | Interview other people in respondents' lives (e.g., partners, children). | Must select roughly 3 years prior to fielding. | Large | Proposals can be sent to NLSY97 PI. Must be done through BLS contract. |
| | Add new respondents (e.g., immigrant refresher sample). | Must select roughly 3 years prior to fielding. | Large to very large | Proposals can be sent to NLSY97 PI. Must be done through BLS contract. |
| | Create new cohorts, e.g. NLSY2010 or Children of the NLSY97. | Roughly 5 years from agreement to data collection. | Very large | Proposals can be sent to NLSY97 PI. Must be done through BLS contract. |

*Transform the NLSY97 into Ready Reference on Marriage and Family*

**Create targeted users' guides on marriage and family topics.** The NLSY97 Users' Guide is an extensive document that not only describes the critical elements of data structure, sampling design, weighting and fielding procedures, but also provides overviews of topical content in the areas of: aptitude, achievement and intelligence scores; education; employment; event history; expectations, attitudes, behaviors, and time use; family background and demographic characteristics; health; income, program participation and assets; marital history and fertility; and training. This breadth of content limits the detail that can be presented on any one set of outcomes. At the same time, variables pertaining to topics such as relationship pathways and fatherhood number in the thousands and cut across topical modules, making it difficult to fully exploit the currently available data.

Targeted users' guides could present the NLSY97 data from the perspective of marriage and family researchers. Guides could be developed, for example, pertaining to fatherhood, step-parenting, teen parenthood, healthy marriage characterization, or a variety of other topics. Content could include a more extensive discussion of relevant variables, a clearer delineation of variables by generation, caveats pertaining to particular measures or subgroups and information about missing values, control totals for key subgroups as well as the software code used to identify them, and cross references to other outcomes of importance, such as employment or health. As one possible first step in this process, a working group could be convened consisting of subject matter and data experts to discuss the topical and technical content.

**Create On-Line Table Generator of Family Variables and Covariates.** BLS provides extensive on-line support with respect to downloading the NLSY data and documentation, which is highly useful for those researchers undertaking the complex analyses that NLSY97 can support. However, some important and highly visible questions in the area of marriage and the family are primarily descriptive. People in the research and policy communities looking for quick and simple tabulations may avoid using the NLSY97 because the effort to understand the data and identify variables of interest is prohibitive for them.

One approach that could add significant currency and visibility to the NLSY97 data would be development of an on-line table generator similar to that provided by the Census Bureau for several of its data sets. Content could include not only outcomes related to marriage and family formation, but also the intersections of these variables with other important domains such as employment, education and health. A possible first step in this process could be the convening of subject matter and data experts to discuss the topical and technical content, as mentioned above.

**Provide Additional Created Variables.** Providing a number of identifying variables, based on existing data, could be particularly helpful in clarifying relationships, such as whether the respondent's cohabiting partner is a biological parent of her child, and the

type of relationship a respondent has with the children's other parent. Suggestions on which new variables to construct could be solicited from experienced NLSY97 users, possibly through the on-line collaboratory or message board suggested below.

**Define Universe of Eligible Respondents by Variable.** On-line documentation for each variable specifies the number of valid and invalid skips in the data, as well as the survey questions leading into and following from that variable. While this provides essential information for assessing the quality of a particular measure, interpreting the meaning of valid skips remains challenging. If, for example, 80 percent of a researcher's sample of nonresident fathers shows a valid skip for the amount of child support paid, this skip may arise either because those nonresident fathers already indicated elsewhere that they paid none, or because the researchers incorrectly identified the sample of nonresident fathers. To know for sure, one must trace through a sometimes lengthy list of preceding variables to untangle skip patterns. This is due to the fact that the pathway of questioning is highly tailored to each respondent's circumstances through a highly complex CAPI instrument.

One enhancement that would make the interpretation of valid skips more transparent would be inclusion of a definitional statement for each variable identifying the universe of eligible respondents. This could be done manually by any knowledgeable data user, with an estimated time of 10 minutes per variable. Given a list of variables totaling over 20,000, a particular user would likely undertake this task only for a variable subset of particular interest. Alternatively, BLS and its contractors could enhance the CAPI program such that this information is produced globally for all variables. While this appears doable in theory, additional work would be needed to assess more closely the feasibility and level of effort needed.

*Create Common Knowledge Base and Platform for Collaboration/Leveraging*

**Create On-Line Collaboratory or Message Board.** Another type of on-line resource that could greatly enhance the usability and visibility of the NLSY97 is a message board, collaboratory or other on-line discussion forum dedicated to uses of the NLSY97 data for research on marriage and family. This forum would enable and encourage researchers to post findings, announce events, discuss methodology, and contribute to the planning of future NLSY enhancements. The forum could also be used to discuss difficult constructs or share knowledge and encourage family research that extends across policy domains. Such a forum could be pursued either through BLS and its contractors, or independently. Either way, one possible first step would be to convene a working group of subject matter and technical experts to discuss needs and structure.

**Create On-Line Repository for Code, Tabulations, Measures, etc.** Similarly, one could create an on-line repository of computer code, tabulations, and definitions of key measures or subpopulations related to NLSY97 research on marriage and the family. Such a repository would make important advances in the field available to a wider spectrum of researchers, and thus facilitate greater progress on important unanswered questions pertaining to relationships and family well-being.

**Promote Analyses of Key Topics Using Available Data.** Analyses on a number of key topics using the data as they currently exist also could be instrumental in making the utility of these data more visible. One such topic suggested by the panel is the extent to which marriage outcomes are driven more by opportunities in the marriage market or personal values related to marriage, and the ramifications of that for programs or policy. Research on these and other topics of interest could be encouraged through the provision of small dissertation grants or other funding opportunities. Such steps could be pursued through a wide range of avenues and organizations.

*Deepen Understanding of Families and Planning for the Next Generation*

**Add Key Survey Questions.** While the NLSY97 data in its current state offers important opportunities to study family relationships and well-being, the panel felt that in order to pursue the important topics on the forefront of the field, new survey questions are needed on topics such as:

- relationship skills
- resource and power sharing
- balancing work and family
- financial literacy
- direct questions on the reasons for not engaging in particular events (e.g., working, getting married, paying child support or visiting nonresident child)
- direct questions on the timing and outcomes associated with key transition events
- more measures of co-parenting and the role of extended family
- in-kind child support provided by nonresident parent of R's child
- perceived exit costs regarding current relationship
- trust of partner or worries of infidelity
- participation in or awareness of healthy marriage services
- occurrences of domestic violence
- the distinction between cohabitors who clearly see themselves on a path toward marriage versus those who do not
- outcomes of the focal respondent's child
- a more extensive rostering of child/parental relationships inside and outside the household.

Changes to the survey can only be pursued through BLS and its contractors, and work would need to begin well in advance of the next wave of data collection. The cost of adding a question depends on the complexity of the question proposed and the amount of development and CAPI programming that would be needed. Since space on the NLSY97 survey is limited, and proposals for new questions would have to be evaluated within that context.

**Interview Additional People in Respondents' Lives.** Panel members felt it was important to interview beyond the focal respondent for two primary reasons. First, they felt that capturing spheres of influence beyond the household is critical to understanding child wellbeing, given the increasing likelihood that children spend significant amounts of time being cared for in other households. Additional interviewees could include

nonresident parents, partners or extended family members. Second, the panel believed that current measures such as relationship quality or parental engagement would be more reliable if gathered from both parties involved. In addition, panel members suggested that some questions could also be asked of the focal respondents' children once they are older. Proposals for such expansions could be sent to the PI of the NLSY97, and would need to be approved by BLS.

**Add New Respondents to Sample.** The panel suggested that the current set of respondents be refreshed with a supplemental immigrant sample to make it more representative of today's young adults. They cited important differences in family formation, employment and other outcomes among immigrants, their growing significance in the population, and the policy relevance of these dynamics. Major changes to the sampling frame can only be pursued through BLS and its contractors, require significant lead time in advance of the next wave of data.

**Create New Cohorts.** The most extensive steps the panel discussed pertain to development of new cohorts within the NLS program. New cohorts could represent children born to mothers in the NLSY97, or an entirely new cross section of youth in 2010. The 2010 option is more likely and, as noted, is already the topic of serious discussions between BLS and other government agencies including the Department of Defense. Organizations interested in the content of future NLSY cohorts should become engaged in such discussions well in advance.

## Conclusions

Proceedings from the ACF-sponsored workshop made it clear that the NLSY97 data can greatly inform key research and policy questions in the area of marriage strengthening and child and family well-being. Several additions or changes could further enhance the utility of the data for these purposes. For example, the data could answer questions directly related to the design of healthy marriage interventions such as:

- What are major barriers to forming and maintaining healthy marriage in society?
- What relationship skills and qualities appear most critical for healthy marriage?
- What are the differences among couples (e.g., by income or race/ethnicity) in types and qualities of relationships and family structures?
- Which couples are at greatest risk and might benefit most from healthy marriage interventions?
- What sources of challenge and resilience among couples appear most relevant for healthy marriage and child well-being?

Similarly, the data could answer important questions that reach across marriage and other domains of the family context and reflect their intersections. For example:

- What can be learned about how some families thrive when facing adverse conditions? What contributes to movement of disadvantaged families up the economic income scale? What role do public programs and social structures play?
- How can we better understand today's complex family structures and identify ways to help ensure that all parents (biological and step, resident and non-resident) are appropriately engaged in the child's wellbeing?
- How can we better understand the relationships between decisions regarding family formation and child bearing and other critical decisions such as those regarding employment opportunities or living arrangements/housing and the implications for each domain?
- What points over a child's life appear to offer the greatest opportunities for families and society to make a difference in their well-being, and at what points are they most vulnerable?

Many of these changes could be accomplished with relatively modest resources in 12 months or less, and would be fairly straightforward to implement either through BLS or independently. These smaller scale enhancements might provide the most likely targets of opportunity in the short run.

The panel suggested, however, that some attention to medium and larger scale opportunities would also be beneficial. Due to the lengthy lead time needed to plan for new content areas or cohorts, planning and design work is needed long before the new data would actually be available. The current discussion around a 2010 NLSY cohort is a clear example of the need for forward planning. Because some of these efforts are costly, they may require more collaboration and leveraging of resources (both funding and expertise) across organizations. Individuals and organizations that could benefit from the content of these more complex changes may wish to build a dialogue around their possible implementation.

Better understanding of how families are doing, what challenges they face, and what helps them thrive is critical within social science research. As a longitudinal data set of respondents just now reaching their early 20s, the NLSY97 is well positioned to illustrate how families are formed and how they grow and function, as well as how this process interacts with elements of work, community, education, faith, and public supports. As many of the young adult respondents in the NLSY97 have already entered parenthood, this is a particularly strategic time to stimulate interest in these data and consider ways to improve their utility for marriage and family researchers.

## Appendix A

The NLSY data are publicly available and readily accessibly, as are the users guide, handbook and codebooks. The primary ways for accessing the NLSY97 data and its documentation are described below:

1. The NLS can be accessed using an on-line search and extraction software known as Web Investigator. This software is on an Internet site where users can search, view, and extract NLS data. A users' manual, also available on-line, shows researchers how to use Web Investigator to access and extract variables of interest.
2. Data sets are available on a CD, using search and extraction software originally developed by CHRR, at the Ohio State University. Because the data sets are zipped on the CD, a single disc can hold data for all NLS cohorts. Users must unzip the cohorts of interest and install the data and search/extraction software on their hard drives. A new all-cohorts disc is pressed each time a new round of data is prepared for any of the active NLS cohorts. Each CD is distributed with a brief hard-copy users' manual; a longer electronic guide to using the search and extraction software is included on the CD.
3. Users may download the CD image from the order/download Internet site. This provides the user with the exact image that is on the CD for that cohort (all individual cohorts are separate downloadable files). The downloaded data are then stored on the user's hard drive, and the search/extraction software is installed on the user's computer as well. The downloaded image also includes a manual explaining use of the search/extraction software.

There is an NLS electronic bibliography accessible on the Internet at <http://www.nlsbibliography.org/>; the bibliography site is fully linked with the nlsinfo.org site for easy user access. The NLS bibliography database houses over 5,200 summaries of ongoing and completed research projects--including published journal articles, government reports, policy analyses, monographs, working papers, conference presentations, doctoral dissertations, and master's theses.

Multiple search indices exist for the on-line NLS Bibliography. These include keywords; author names (including truncation); citation title; journal names; format of publication; cohort; year of publication; abstract; and full text. The list of formats includes full text on-line, in-progress research, journal article, master's thesis, monograph or chapter, newspaper, conference presentation, Ph.D. dissertation, report, and working paper. An 'advanced search' strategy allows the user to integrate up to seven of these search types in a focused and targeted manner. The recently added abstract search option allows users to enter any words and search for abstracts containing these terms. This search is useful when the supplied term does not exist as a keyword or has not been assigned to a citation. These indices are offered on-line to the user as search types that provide access to the annotated bibliography.