

## 6. PROSPECTS FOR ENHANCING FEDERAL SURVEYS

### 6.1 Introduction

As noted in previous chapters, the current literature provides an extensive understanding of the characteristics and service needs of currently homeless families, yet there remain substantial knowledge gaps that make it difficult to develop an accurate and useful typology of homeless families. These gaps include the following:

- Data on homeless families across various regions of the country;
- Data on key subgroups, such as:
  - Families at risk of becoming homeless;
  - Moderate need homeless families;
  - Families that fall back into homelessness despite intervention;
  - Working homeless families; and
  - Two-parent homeless families.
- Longitudinal studies of homeless families; and
- More intensive studies of homeless children.

It was noted in Chapter 4 that none of the general population studies currently or recently conducted by the Federal Government, such as the Current Population Survey (CPS), the Panel Study of Income Dynamics (PSID), or the National Longitudinal Surveys of Labor Market Experience (NLS) can address these knowledge gaps in their present form.

Given the size, scope, and resources already invested in conducting various national surveys, it would be useful to determine if there are surveys that are ongoing, or planned for the future, that might potentially be enhanced to fill these gaps. In this chapter, current and planned survey efforts are examined and three surveys are identified that could be enhanced to provide useful information on families who have experienced homelessness one or more times, and families who are at risk of homelessness. A short battery of questions is proposed that could be added to each identified survey to strengthen the ability of each to address one or more of the gaps in the knowledge and understanding of homeless families.

## **6.2 Overview of National Survey Efforts**

A number of national surveys are regularly conducted to address a myriad of information needs. These surveys are generally sponsored, if not actually conducted, by the Federal Government, from the basic census task of describing how many people live in the country in order to apportion congressional seats and Federal spending, to more focused efforts designed to provide both private and public officials with timely, reliable, and accessible information on such topics as labor force participation and income, housing, and health and nutrition.<sup>7</sup> In general, these survey efforts can be divided into three broad types:

1. Ongoing cross-sectional studies;
2. Short-term longitudinal studies; and
3. Long-term longitudinal studies.

Each of these survey types provides a different set of opportunities and challenges with respect to the information it can already provide on families that are at risk and/or have experienced homelessness, as well as for its potential to be enhanced to provide such information.

## **6.3 Review of Cross-Sectional Surveys**

The national cross-sectional surveys currently in operation are designed to provide current information on various topics (e.g., the percentage of the population currently working, health status of people, or the extent of illegal substance abuse). These surveys typically collect information on a large number of people in order to be able to provide accurate and reliable estimates not only at the national level, but also for smaller geographic subunits, such as the state, metropolitan region, city, or even census tract level.

---

<sup>7</sup> The surveys examined in this chapter were identified using a variety of sources. In addition to the surveys identified and examined in Chapter 2 and recommendations made by members of an Expert Panel brought together in July 2005 [see Chapter 3], surveys were identified through various web searches. Summaries and lists of databases, such as the list of public databases maintained by the American Sociological Association were also reviewed. Two recent government reports were also reviewed that discussed similar recent efforts at examining various Federal surveys to make more efficient use of these data collection sources. One was an inventory of Federal databases conducted for the HHS Office of the Assistant Secretary for Planning and Evaluation, as part of an assessment of major Federal databases for analyses of Latinos and Asian or Pacific Islander subgroups and Native Americans (Waksberg, Levine, and Marker, 2000). The second was a more recent review of Federal health surveys sponsored by The Robert Wood Johnson Foundation and the CDC's National Center for Health Statistics (AcademyHealth 2004).

In terms of providing useful information on families that are, have been, or may be homeless in the future, these cross-sectional, general population surveys have several major advantages:

- The ability to understand factors that helped families exit homelessness;
- Depending on the size and structure of the data set, the ability to examine at-risk and literal homelessness for subgroups of families, including:
  - Working poor families;
  - Moderate-need poor families; and
  - Two-parent poor families.
- The ability to develop estimates (albeit, likely underestimates) of the incidence and prevalence of homelessness among families over a specific period of time at the national level and, depending on the size and structure of the data set, at the regional and/or state level, and the ability to examine change in the incidence, prevalence, and characteristics of homeless families over time.

Cross-sectional studies also have two major limitations for use in the current effort. First, depending on the sampling frame and data collection methods used, a study may exclude currently homeless families. A study that recruits participants from a list of addresses that includes only homes, apartments, and condominiums, for example, would exclude not only those who are living on the streets, but those living in emergency shelters and other types of temporary housing. Likewise, a survey that collects information only by phone could not include people who do not have their own phone, which is likely to be true for most homeless families (as well as families at risk of becoming homeless). As a result, these studies would provide an underestimate of the overall incidence and prevalence of homelessness. Second, these studies can only examine past homelessness, with no opportunity to examine families prospectively. These surveys generally offer large samples, but either select different samples each time data are collected or do not provide the ability to link responses across different collection points.

Table 6-1 presents a summary of the nine major, cross sectional surveys that were identified and reviewed for this effort. Each survey is described according to the type of sampling frame used (i.e., how the sample was initially drawn or identified), the size and composition of the sample (i.e., if the data are collected on individuals, households, or both), the frequency of data collection, whether the sampling frame is supplemented by a specific oversample (e.g., oversample of low-income households),

**Table 6-1. Overview of Federal cross-sectional survey efforts**

Survey	Sampling frame	Sample size and type	Frequency	Oversamples	How data collected	Primary focus	Other notes
American Community Survey (ACS) (Conducted by Census Bureau)	National area probability	800,000 households	Annually	None	Mail (50%)	Demographic Housing Social Economic	ACS replaces the decennial census long form
	Currently excludes group quarters, expected to include in 2006	3 million households starting in 2005  Data collected on all household members			Computer-assisted telephone surveys (CATI)  In-person (sample of nonresponders)		
American Housing Survey (AHS) (Conducted by Census Bureau)	National area probability (excludes group quarters)	55,000 households – national survey	Biannually – national survey	None	CATI In-person	Size, composition, and state of housing stock	Survey returns to the same address for each wave, even if the household has changed
	Metropolitan area probability surveys collected as well	3,200 households – for each metropolitan survey	Every 6 years for each metropolitan survey, conducting 14 per year				

**Table 6-1. Overview of Federal cross-sectional survey efforts (continued)**

Survey	Sampling frame	Sample size and type	Frequency	Oversamples	How data collected	Primary focus	Other notes
Current Population Survey (CPS) (Conducted by Census Bureau)	National area probability	60,000 households 130,000+ people  Data collected on all household members	Monthly  Households in survey for 4 months, out 8, in 4, and then dropped	Latinos (March sample of each year)	Initial Interview In-person  In-person or CATI for followups	Labor force participation	Supplemental questions regularly added: - March: Annual demographic survey - Housing vacancy survey
National Health and Nutrition Examination Survey (NHANES) (Conducted by Nat. Center for Health Statistics)	National area probability	5,000 people	Annually	Low-income Whites Adolescents Persons 60+ Blacks and Latinos	In-person  Additional medical exams at a mobile exam center	Health Nutrition	Data combined and released in 2-year waves
National Health Interview Survey (NHIS) (Sponsored by Nat. Center for Health Statistics)	National area probability (includes group quarters)	43,000 households 106,000 people  Data collected on all household members	Annually	Blacks and Latinos	In-person	Health and illness Disability	Topical supplemental modules regularly included

**Table 6-1. Overview of Federal cross-sectional survey efforts (continued)**

Survey	Sampling frame	Sample size and type	Frequency	Oversamples	How data collected	Primary focus	Other notes
National Household Education Survey (NHES) (Conducted by National Center for Education Statistics)	National RDD	‘2003 – 32,000 Households  Limited household data, more on selected adults and children	Biannually	Blacks and Latinos	CATI	Various educational activities of adults and/or children	
National Immunization Survey (NIS) (Sponsored by National Center for Health Statistics)	National RDD  Screen 1 million households to find families with children 19 to 35 months	35,000 households 94,000+ people  Data collected on family, sample adult, and sample child, if available	Annually	None	CATI	Immunization	
National Survey of America’s Families (NSAF) (Conducted by Urban Institute)	National RDD supplemented with area probability in poorer neighborhoods	Three cohorts: 1997 – 45,000 households 1999 – 46,000 households 2002 – 40,000 households  Data collected on adults and one child if available	Three separate cohorts  No future surveys scheduled at this time	Oversampled in 13 large states  Low-income	CATI (majority)  In-person for households w/o phones	Employment Education Social services Financial services	

**Table 6-1. Overview of Federal cross-sectional survey efforts (continued)**

Survey	Sampling frame	Sample size and type	Frequency	Oversamples	How data collected	Primary focus	Other notes
National Survey on Drug Use and Health (NSDUH) (formerly National Household Survey on Drug Abuse) (Sponsored by SAMHSA)	Area probability sample by state (to provide valid state estimates)  Includes group quarters (e.g., shelters, rooming houses)	70,000 people  Randomly selected persons per household	Annually	Not currently (earlier oversampling of Blacks and Latinos stopped when the sample size was increased)	In-person (including audio computer assisted self-interviewing ACASI)	Cigarette use Illicit drug use Alcohol use Mental illness Mental health treatment	NSDUH notes that the sample sizes for group quarters are too small to provide valid estimates

how the data were collected (e.g., in person, by telephone, or some combination), the primary content focus of the survey, and any other notes that help us understand the suitability of the survey for informing the typology.

These features were examined to identify surveys that offer the best opportunity to be enhanced to inform efforts to develop a typology of homeless families. Four criteria were used to select candidates for enhancement:

- Whether the survey is still being conducted;
- Whether the sample design (frame, size, type, and frequency) and data collection methods are more likely to include recently homeless families, as well as currently unstable families;
- Whether the data are collected on family characteristics; and
- Whether the sample size is large enough to examine subpopulations, regional, and state differences in homeless families, and families who are doubled-up.

Only two studies, the Current Population Survey (CPS) and the American Community Survey (ACS), met all four of these criteria. In this section, the rationale for eliminating the seven other studies from further review is explained and then the opportunities offered by the CPS and ACS surveys are described in more detail.

### **6.3.1 Studies No Longer Being Conducted**

The National Survey of America's Families was eliminated from further consideration as there are no current plans for extending its data collection to a fourth cohort of respondents. It has a number of features that would have made it a good candidate for enhancement, including an oversampling of poorer neighborhoods, relatively large samples, and a focus on a number of data elements that could be fruitfully used to address questions about homeless families, including employment, education, and social service use. If a new NSAF study is mounted, however, it might be useful to consider including questions about homelessness.



### **6.3.2 Study Design and Structure Likely to Exclude Recent Homeless or Residentially Unstable Families**

Three studies—the National Immunization Survey (NIS), National Household Education Survey (NHES), and American Housing Survey (AHS)—are not good candidates for enhancement because they use sample designs and/or data collection methods that are likely to exclude current and recently homeless families, as well as families that are currently residentially unstable. The NIS and NHES surveys use random-digit dialing (RDD) to identify study participants. Random digit dialing involves selecting telephone numbers at random from a frame of all possible telephone numbers. While RDD is a reliable and efficient method for randomly selecting a national sample, unless a currently homeless person or family happens to have a cell phone, RDD will exclude people and families who are currently living on the streets and/or in shelters. It is also likely to undersample those who are precariously housed, since they are likely to be part of the small percentage of households that do not have a phone or have phone numbers that are routinely disconnected.

In addition to these problems with their sampling frames, the NIS and NHES use computer-assisted telephone surveys (CATI) to collect data. Reliance on the telephone to collect data is further likely to lead to an underreporting of both current and recently homeless people and families. The AHS uses an area probability sample to identify study participants. In this approach, the country is broken down into various geographic units, with the smallest often having only 100 to 200 housing units (e.g., street addresses), and various methods are then used to randomly select these small geographic units or segments. Area probability samples have a better chance of including homeless families in their data sample, however, when they include group quarters, such as homeless shelters and transitional housing, in their sample frame. Unfortunately, the AHS excludes group quarters from its sample design. Furthermore, the AHS design of interviewing households living at the same address initially selected (e.g., returning to 100 Main Street each time), even if the household living there has changed since the previous survey, minimizes the likelihood of identifying homeless and at-risk families.

### **6.3.3 Family Data Not Collected**

Two studies, the National Health and Nutrition Exam (NHANES) and the National Survey on Drug Use and Health (NSDUH), were dropped from consideration because both collect data mainly on a specific individual rather than a family or household. This is a particularly unfortunate feature, since the

NSDUH annually collects data on a large number of people (more than 70,000), with samples designed to provide valid estimates at the state level and using a sample frame that includes group quarters. Furthermore, the NSDUH collects information on a number of domains that might be useful to examine in relationship to both prior homelessness and the risk of homelessness, including illicit drug use, alcohol use, mental health status, and mental health treatment.

#### **6.3.4 Studies Unable to Examine Subpopulations or Regional/State Differences**

The National Health Interview Survey (NHIS) uses a national area probability sample; collects information on health, illness, and disability that could be usefully examined in relationship to literal and at-risk homelessness; and regularly includes supplemental questions. The major challenge with the NHIS is its sample size. With a total sample of 43,000 and using a 1.5 percent yearly incidence rate of family homelessness (Burt et al., 1999), the NHIS would likely produce 600 to 700 cases per cohort and would not provide the ability to examine specific subgroups or data on homelessness at any level other than national.

#### **6.3.5 Studies that Met Primary Selection Criteria**

As noted earlier, only two studies meet all four of the primary selection criteria: the Current Population Survey and the American Community Survey.

**Current Population Survey.** The CPS is the main source of labor statistics in the United States. Conducted monthly by the Census Bureau for the U.S. Bureau of Labor Statistics, the CPS typically interviews a nationally representative sample of approximately 50,000 households. Respondents are selected using a national area probability sample. Part of the sample is changed each month; that is, a selected household or address is in the sample for 4 months, taken out for 8 months, put back in for 4 months, and then entirely removed. Given this rotation process, three-fourths of the sample stays the same from one month to the next, and half of the sample is surveyed in the same month from one year to the next. The monthly responses are not linked, however.

The CPS collects information on each member of the selected household aged 15 or older (although published reports focus on people ages 16 or over). Information collected includes data on

employment, hours of work, and income, in addition to such demographic characteristics as age, sex, race, marital status, and educational attainment. Supplemental questions are also frequently included with the CPS. The results from each March survey, for example, are used to develop the Annual Demographic Supplement for the U.S. Census. In order to provide an adequate sample to do in-depth analyses of the Latino population, additional Latino sample units are added to the survey in this month.

With approximately 50,000 households selected each month, the CPS provides an opportunity to identify families that have been recently homeless or are at risk of becoming homeless. The broad geographic spread of the survey could help determine rates of homelessness across various regions of the country, as well as differences among urban, suburban, and rural areas. Information obtained over time could also be used to monitor changes in the percentage of families/individuals that have been homeless. In order to provide this sort of information, though, questions would need to be added about recent homeless and housing experiences.

**American Community Survey.** The ACS is a new survey effort being conducted by the Census Bureau and is designed to replace the long form of the decennial census. The main reason for this change is that the information provided by the long form tends to be increasingly out of date later in the decade. The ACS will enable the Census Bureau to provide more frequently updated information on the same range of topics that are covered in the decennial census.

Respondents for the ACS will be selected using a national area probability sample. Since the ACS is still being field tested, the survey initially included only 800,000 households, and group quarters were excluded from the sample. By 2006, however, group quarters, including emergency homeless shelters, transitional shelters, temporary housing, and hotels or motels used to provide housing for people without conventional shelter, were to be included.<sup>8</sup>

The ACS is designed to collect the same information as the long form, such as demographic, housing, social, and economic data. Information is obtained on every person in the household. Data for the ACS will be collected using three data collection methods. The first step will be self-administered mail surveys; it is expected that at least half of the responses will be obtained this way. Households that have not responded by mail will then be contacted by telephone. Finally, attempts will be made to conduct in-person interviews with at least a sample of those still remaining.

---

<sup>8</sup> In order to protect the confidentiality of their locations, group quarters will not include domestic violence shelters.

When it is fully operational, the ACS is expected to collect information on over three million households annually, making it by far the largest survey effort in the country. The sample size of the ACS should be large enough to provide valid annual estimates for every state, as well as all cities, counties, and metropolitan areas with 65,000 people or more. For smaller areas, such as rural areas or individual census tracts, results will have to be aggregated over a 3- to 5-year period to produce a sufficiently large sample.

**Prospects for Survey Enhancement.** Of the eight national cross-sectional surveys examined and summarized in Table 6-2, only the CPS and the ACS offer benefits for obtaining information on at-risk and literally homeless families. Of these two surveys, the ACS is the more useful for several reasons. First, the ACS has a much larger sample than the CPS. Questions about homelessness and the risk of homelessness added to the ACS would be asked to over three million households annually, while supplemental questions to the CPS would likely be asked only one month a year, to a sample of 50,000 households. Second, the data collection methods used for the ACS are more likely to locate and include precariously housed families, as the survey will eventually include families living in emergency homeless shelters and temporary housing. The data collection procedures used by the CPS provide much less opportunity to locate people who cannot be contacted initially. Finally, the CPS collects a relatively small amount of information compared to the ACS, with a major emphasis on labor force participation that is likely to be less useful in developing a typology of homeless families.

Given these additional considerations, the ACS offers the best prospects for addressing knowledge gaps about homeless families, if it is enhanced. Given its large sample size of over 3 million households a year, for example, the ACS could provide an opportunity to look at homelessness in specific geographic areas, providing an ability to examine how market forces, social capital, and other contextual variables relate to the incidence of family homelessness.

**Table 6-2. Cross-sectional surveys that meet selection criteria for possible enhancement**

Selection criteria	American Community Survey	American Housing Survey	Current Population Survey	National Health and Nutrition Survey	National Health Interview Survey	National Household Education Survey	National Survey of America's Families	National Survey on Drug Use and Health
Surveys still being conducted	✓	✓	✓	✓	✓	✓	No	✓
Sample design and data collection methods less likely to exclude recently homeless and currently unstable families	✓	No	✓	✓	✓	No	✓	✓
Data collected on family characteristics	✓	✓	✓	No	✓	✓	✓	No
Sufficient sample size to examine: - Subpopulations - Regional/state differences - Doubled-up	✓	✓	✓	No	No	✓	✓	✓
Candidate for enhancement?	✓	No	✓	No	No	No	No	No

The sampling frame for the ACS already plans to begin to include overnight shelters and other facilities where homeless families could be found. Even if the ACS sample includes only a percentage of families found in the nontraditional housing settings, its large sample should still yield a large absolute number of homeless families that could be examined. Again, using a yearly incidence rate of family homelessness of 1.5 percent (Burt et al., 1999), the ACS could produce a sample of 45,000 homeless households a year. Even at half that rate, there would still be 20,000 to 25,000 homeless households in the sample. Furthermore, because the ACS is still being developed and refined, it may be possible to refine the sampling procedures to better ensure that emergency and transitional shelter facilities that serve homeless families and individuals are part of the sample frame.

#### **6.4 Review of Longitudinal Studies**

In addition to the cross-sectional surveys, there are several longitudinal studies that track the same person, family, or household over time. Because of the challenges and costs involved in tracking respondents, these surveys typically involve much smaller samples than cross-sectional studies and are often much more focused on specific populations and/or topics. Some of these surveys, such as the Survey of Income and Program Participation (SIPP) or the Medical Expenditure Panel Survey (MEPS), track people or households for only a few years. There are also two well-known, long-term longitudinal studies to consider: the Panel Study of Income Dynamics (PSID), begun in 1960, and the National Longitudinal Survey of Labor Market Experiences (NLS).

Longitudinal studies offer many of the same potential advantages as cross-sectional studies, and they have the added potential benefit of tracking people over time and thus may provide an opportunity to examine entries into, and exits out of, homelessness (depending on their tracking methods). However, as discussed in a later section, the longitudinal studies are smaller in overall sample size and lack the ability offered by cross-sectional studies to examine regional differences as well as various subgroups.

Table 6-3 presents a summary of the eight longitudinal surveys that were identified and reviewed for this effort according to the same features used to review the cross-sectional surveys.<sup>9</sup> As with the cross-sectional surveys, the longitudinal surveys were initially examined according to four key selection criteria to identify surveys that offer the best opportunity to be enhanced to inform efforts to develop a typology of homeless families. Six of the eight surveys were deemed inappropriate candidates for enhancement, as discussed later. Only two surveys—the NLS 1979 cohort study and the NLS 1997 cohort study—met all four of the initial criteria.

#### **6.4.1 Studies No Longer Being Conducted**

Two of the longitudinal studies described in Table 4-3 were not considered appropriate candidates for enhancement, either because they have just finished or will soon end data collection. These include both the kindergarten and birth cohort samples of the Early Childhood Longitudinal Study (ECLS) and the National Survey of Child and Adolescent Well-Being (NSCAW). The ECLS kindergarten cohort ended data collection in 2004, while the birth cohort is expected to end data collection in 2008. The NSCAW study ended late in 2005.

#### **6.4.2 Study Design and Structure Likely to Exclude Recent Homeless Families or Residentially Unstable Families**

The MEPS was not considered a good candidate for enhancement because it uses a sample design that appears to make it more difficult to include recently homeless families as well as families that are currently at risk of being homeless. While most of the longitudinal studies use some sort of national area probability sample to select their respondents, the MEPS sample is selected from households identified through the NIS, which in turn identifies families using RDD. As previously discussed, it is expected that the use of RDD to identify study participants will further reduce the likelihood of a study including currently homeless people or families, and those who have been recently homeless or who are residentially unstable.

---

<sup>9</sup> Since the NLS79 and NLS97 studies collect information on two distinct cohorts of households, and even use different data collection instruments, they are treated as two separate studies for the purposes of this review.

**Table 6-3. Overview of Federal longitudinal surveys**

Survey	Sampling frame	Sample size and type	Frequency	Oversamples	How data collected	Primary focus	Other notes
Early Childhood Longitudinal Study – Birth Cohort (ECLS – Birth) (Sponsored by National Center for Education Statistics)	National random sample of birth certificates (or hospital records)	13,500 children born in 2001	Five waves of data collection: - 9 month - 18 month - 4 years - Kindergarten - 1st grade  Data collection ending in 2008	Asian, Pacific Islander, Chinese  Low and moderately low birthweight  Twins	In-home interviews with parent/ guardian  1-1 child assessments	Child and family characteristics that influence school preparedness	
Early Childhood Longitudinal Study – Kindergarten Sample (ECLS – K) (Sponsored by National Center for Education Statistics)	National area probability of elementary schools	22,000 children in kindergarten 1998-99  Information collected on/from: -children -parents -teachers/school administrators	Most data collected annually for 6 years (K-5th grade)  Some data collected semi-annually for first 2 years  Data collection ended in 2004	None	Various methods: - 1-1 assessment - child interviews - CATI (parents) - self-administered (teachers, administrators)	Impact of early and middle-childhood education	



**Table 6-3. Overview of Federal longitudinal surveys (continued)**

Survey	Sampling frame	Sample size and type	Frequency	Oversamples	How data collected	Primary focus	Other notes
Medical Expenditure Panel Survey (MEPS) (Sponsored by Agency for Healthcare Research and Quality)	National area probability  (Based on NIS RDD sample)	7,000 – 13,000 households  New waves added annually	Five interviews conducted over 2 years	Blacks and Latinos  Low income  Elderly	In-person CATI	Health care use and expenditures	
National Survey of Child and Adolescent Well-Being (NSCAW) (Sponsored by Administration for Children and Families)	Children in welfare agencies nationwide (97 different agencies)	5,400 children  700 supplemental sample	Three to four waves of data collection: - baseline - 12 month - 18 month - 36 month possible  Project ending in 2005	Supplemental sample (700 children) in foster care	In-person	Demographic characteristics of children and families  Pathways and services utilized	
Survey of Income and Program Participation (SIPP) (Conducted by Census Bureau)	National area probability	2001 cohort – 36,700 households  (Only original sample members reinterviewed)	Every 4 months over 3 to 4 years  2001 cohort just ended	Low-income	In-person CATI	Labor force Income Program Participation and eligibility	

**Table 6-3. Overview of Federal longitudinal surveys (continued)**

Survey	Sampling frame	Sample size and type	Frequency	Oversamples	How data collected	Primary focus	Other notes
National Longitudinal Surveys of Youth 1979 (NLSY79) (Sponsored by Bureau of Labor Statistics)	National area probability sample youth/young adults  Initial NLS samples started in 1968, ended 1981  New cohorts added in 1979 and 1997	12,686 youth ages 14 to 22 in 1979  7,724 respondents in 2002 sample	Annually 1979-94  Biennially starting in 1994	Latino, Black, and economically disadvantaged nonminority  Young adults in the military (discontinued in 1985)	Initially in-person  Mostly CATI in recent years	Labor market activities	Supplemental questions have been added at various waves
National Longitudinal Surveys of Youth 1997 (NLSY97) (Sponsored by Bureau of Labor Statistics)	National area probability youth/young adults	8,984 youth ages 12 to 17 in 1997	Annually, 1997-2003	Black or Latino youth	Initially in-person  CAPI ACASI	Education Labor market behavior Family and community Background	Supplemental questions have been added at various waves

**Table 6-3. Overview of Federal longitudinal surveys (continued)**

Survey	Sampling frame	Sample size and type	Frequency	Oversamples	How data collected	Primary focus	Other notes
Panel Study of Income Dynamics (PSID) (Conducted by University of Michigan)	National area probability sample  Supplemental sample of low-income families	4,800 households 65,000+ people	Annually 1968-97  Biennially starting in 1999	Initial supplemental sample of low-income families  Refresher sample added in 1997	Initially in-person  Mostly CATI more recent years (97%)	Income Labor force Marital status	Supplemental questions have been added at various waves

### **6.4.3 Studies Unable to Examine Subpopulations or Regional/State Differences**

Two longitudinal studies, the SIPP and the PSID, that in many respects appeared to be good candidates for enhancement, were eventually considered to have samples that were too small to provide reliable estimates of recently homeless or residentially unstable families.

The SIPP is a series of national panel studies designed to collect information on income, labor force participation, and participation and eligibility for various government programs. The length of time each panel is followed has varied in recent years, from 2.5 to 4 years. Sample sizes have also varied from cohort to cohort within the panel studies, from 14,000 to 36,700 households in the 2001 study. Even at its largest, however, the SIPP study is likely to identify only 500 or 600 recently homeless families at most (based on a 1.5% annual homeless rate). Although this would be a sufficiently large sample to examine national trends, it would not provide a large enough sample to reliably examine any regional or geographic differences in homelessness. Combined with the fact that the SIPP tracks families for only a few years, it does not appear to be a good candidate for enhancement.

Initially, the PSID offered the best prospects for informing national efforts toward homeless prevention and resource allocation. Begun in 1968 and conducted by the University of Michigan, the PSID originally consisted of two independent samples—a cross-sectional national sample of approximately 3,000 families and a national sample of 2,000 low-income families. From 1968 to 1996, individuals from these initial samples were interviewed annually, including people who may no longer have been living in the original sampled household (e.g., children of the originally selected households). Because it tracked everyone associated with the originally sampled household, by 1996 the PSID had grown to over 65,000 individuals. In order to keep the sample more manageable, as well as to readjust the sample to better reflect the U.S. population, adjustments were made to the sample in 1997 that reduced the number of “core” families and added a new sample of families, particularly Latino and Asian households. The distinct advantages of the PSID with respect to being able to address knowledge gaps about homeless families are the following:

- Longitudinal, currently conducted every 2 years;
- Long history, starting in 1968;
- An oversample of low-income households, who have a higher probability of having been or becoming homeless than the general population;

- A residential followback as part of its data collection, so many of the changes could be adding questions to that part of the instrument; and
- A wealth of data that have been consistently collected over time, such as income sources and amounts, employment, family composition, and demographic changes.

The major limitation of the PSID, however, is its sample size. If 1.5 percent of the households in the current PSID sample experienced homelessness in any given year, this would produce a sample of only 75 families to examine given the current overall sample of 4,800 households. A further complication with the PSID, or with any longitudinal study, is the ability to track and maintain contact with more difficult-to-reach study participants, such as people or families who become homeless. The response rates for the PSID have generally been very high, averaging 97 percent to 98 percent a year.<sup>10</sup> As is noted in the PSID guide, though, even small rates of attrition from wave to wave can create problems over time. In 1988, for example, the response rate for individuals who lived in 1968 households was only 56 percent. Furthermore, the PSID **does not** make an attempt to recontact households that drop out, so even a small level of attrition may severely impact the likelihood of identifying families that have been or become homeless. Thus, despite its many potential advantages, concerns over sample size and composition make the PSID a less than ideal candidate for enhancement.

The only two longitudinal surveys that do seem to have some potential for addressing knowledge gaps about homeless families are the two recent NLS cohorts: the National Longitudinal Surveys of Youth 1979 (NLSY79) and the National Longitudinal Surveys of Youth 1997 (NLSY97). Both the NLSY79 and the NLSY97 are part of the National Longitudinal Surveys conducted for the U.S. Department of Labor, BLS.

#### **6.4.4 Studies that Met Primary Selection Criteria**

**National Longitudinal Surveys of Youth 1979.** The NLSY79 is a series of surveys with a nationally representative sample of 12,686 young men and women who were between the ages of 14 and 22 in 1979. Annual interviews were conducted from 1979 until 1994; since then, respondents have been interviewed every other year (1996, 1998, etc.).

---

<sup>10</sup>These numbers are from the PSID Guide available online at: <http://psidonline.isr.umich.edu/Guide/ug/chap5.html>.

Respondents were selected using a multistage, stratified national area probability sample of dwelling units and group quarters. Three independent probability samples were recruited:

- Cross-sectional sample of 6,111 people designed to be representative of the young adult population living in the United States at that time;
- Supplemental set of 5,295 people designed to oversample Latino, Black, and economically disadvantaged, non-Latino, non-Black youth; and
- A military sample of 1,280 people designed to represent the population born between January 1, 1957 and December 31, 1961, serving in the military as of September 30, 1978. Interviewing of the full military sample stopped in 1985.

Data for the NLSY79 have usually been collected using personal interviews, but telephone interviews have also been used and, in fact, are becoming more common. The NLS studies are primarily designed to study the transition of young people into the labor market. As a result, questions are typically asked about education, work, and training. Information is also collected on everyone living in the household of the initial respondent.

New topics have been frequently added to the NLS surveys. The National Institute on Alcohol Abuse and Alcoholism, together with the National Institute on Drug Abuse, for example, has added questions on alcohol and substance abuse on various NLS waves, while the National Institute of Education added a set of time-use questions to the 1981 survey.

In 1986, the NLS79 was further enhanced with a survey of children from the NLSY79 sample, funded by the National Institute of Child Health and Human Development (NICHD), along with a number of other government agencies and private foundations. These supplemental questions have collected information on the development of children born to NLSY79 women and, starting in 1994, a separate survey was administered to children age 15 or older.

**National Longitudinal Surveys of Youth 1997.** With the aging of the NLSY79 sample, a new cohort of young adults was selected in 1997 to participate in the NLSY97 survey. The NLSY97 sample consists of two independent national probability samples:

- Cross-sectional sample of 6,748 people between the ages of 12 and 17 in 1997 designed to be representative of the young adult population living in the United States at that time and

- Supplemental set of 2,236 people designed to oversample Latino and Black respondents.

Data are usually collected using in-person interviewers although, as with the NLSY79 study, telephone interviews are also conducted and are becoming more common over time.<sup>11</sup> While much of the interview is conducted using a computer-assisted personal interview (CAPI) system, questions about particularly sensitive issues are asked using an audio computer-assisted self-interview (ACASI) procedure. While respondents are living with their parents or other legal guardian, many of the household questions are asked directly to the parents. When the initial respondent is living elsewhere, information is collected on everyone in the respondent's household. Followup surveys are conducted annually, although the gap between the initial survey and the second round turned out to be a little longer, approximately 18 months.

As with all of the NLS surveys, the primary purpose of the NLSY97 is to collect information on labor force experience, education, and the transition into the labor market. A number of additional questions have also been added, however, including a set of questions on crime and criminal activities sponsored by the U.S. Department of Justice, as well as development questions added by NICHD.

**Prospects for Survey Enhancement.** Although several longitudinal studies were initially thought to be able to provide information on knowledge gaps on homeless families, at least if they were enhanced, this review suggests that only the two latest NLS surveys—NLSY79 and NLSY97—may be particularly good candidates. Of these two, the NLSY97 may offer the better opportunity. A major challenge with the NLSY79 cohort is that the primary respondents are moving out of the age when homelessness seems to be most likely to occur. As noted in Chapter 1, the risk of becoming homeless seems to be higher when people are in their mid to late-20s. Therefore, the NLSY79 sample would have been most likely to have experienced homelessness from the mid-1980s to early 1990s. By now, with the youngest members of the NLSY79 sample already 40 years old, this cohort may be too old to provide a good opportunity to examine homelessness, at least prospectively.

The NLSY79 sample does include a subsample of children born to initial study participants whose ages would make them more likely to be currently experiencing homelessness. Adding questions to the NLSY79 sample about their history of homelessness, as well as to the NLSY79 Children and Young Adult surveys about both their history and current incidence of homelessness would, therefore,

---

<sup>11</sup>Only 3 percent of the initial NLSY97 interviews were done over the telephone, for example, compared to 8.7 percent of the interviews in 2000.

provide a rare opportunity to examine the intergenerational effects and impact of homelessness. However, the smaller sample size of the children's sample (only children born to women in the NLSY79 sample are surveyed) makes this a less promising approach.

The NLSY97 sample provides the best opportunity to examine family homelessness prospectively, which could help answer questions about the factors that lead to people becoming homeless and factors that help predict exiting out of homelessness. The attrition rate for the NLSY97 sample has so far been fairly low, making it more likely to still include respondents whose families have been homeless or who are at risk of becoming homeless. For example, as of the last reported round of the NLSY97 surveys (Round 5), 88 percent of the initial sample had been interviewed. The primary reason for not conducting an interview has generally been because the respondent refused the interview rather than an inability to locate the respondent (65% of the nonresponses in Round 5 resulted from refusals).

## **6.5 Proposed Housing Questions**

As noted in Chapter 2, existing studies, including the NLS and ACS, do not provide enough information to identify families that are currently or have recently been homeless. The major enhancements that these surveys need include adding questions and/or adding response categories that make it possible to identify homeless families.

**American Community Survey.** Enhancements of the ACS would focus on the housing section of the survey. These enhancements would include questions to determine whether household members are currently living in some sort of emergency or transitional housing, and whether they have been homeless or at risk of being homeless in the past 12 months. Also proposed is a question on whether anyone in the household has a housing subsidy. Table 6-4 shows in which sections of the survey instrument those enhancements could be made.



**Table 6-4. Possible enhancements to the American Community Survey**

Current Living Situation:	<p>Possibly add after Question 1 in the Housing section:</p> <ul style="list-style-type: none"> <li>■ Are you currently living in an emergency or transitional housing unit or in some other sort of temporary housing? Y/N</li> </ul>
Recent Homelessness or Risk of Homelessness:*	<p>Possibly add at the end of the Housing section: In the past 12 months:</p> <ul style="list-style-type: none"> <li>■ Did you ever not pay the full amount of rent or mortgage payments? Y/N</li> <li>■ Were you ever evicted from your home or apartment for not paying the rent or mortgage? Y/N</li> <li>■ Did you move in with other people even for a little while? Y/N</li> <li>■ Did you stay at a shelter, in an abandoned building, an automobile or any other place not meant for regular housing, for even one night? Y/N</li> </ul>
Housing Subsidy:	<p>Possibly add to Question 15 in the Housing section, which currently asks about food stamps:</p> <ul style="list-style-type: none"> <li>■ At any time in the past 12 months did anyone in the household receive a housing subsidy? Y/N</li> </ul>

\* These are modified versions of questions asked in the Fragile Family Study.

**National Longitudinal Surveys of Youth 1997.** The NLSY97 already collects housing and mobility information. In fact, the NLSY97 uses a set of responses to describe the respondent’s current living situation that already includes “Shelter (for homeless or abused) or on street...” It then follows up with a question concerning how long the person has been living in this place. The NLS97 also includes a number of questions about various risk and protective factors. Many of these questions, including such topics as illegal drug use, criminal behavior, and arrests, are asked as part of an ACASI section.

What is not collected in the NLS97 survey is whether the respondent was homeless at some point between the current and previous interviews for those who moved, and whether the respondent was ever at risk of being homeless. Finally, depending on the length of time it takes to add any of these questions into the NLS, it may also be necessary to include homeless history questions at least once. Table 6-5 shows possible enhancements in the NLSY97.

**Table 6-5. Possible enhancements to the National Longitudinal Surveys of Youth 1997**

<p>Recent Homelessness or Risk of Homelessness.*</p> <p>History of Homelessness:</p>	<p>Possibly add in the Household Information section: In the past 12 months:</p> <ul style="list-style-type: none"><li>■ Did you ever not pay the full amount of rent or mortgage payments? Y/N</li><li>■ Were you ever evicted from your home or apartment for not paying the rent or mortgage? Y/N</li><li>■ Did you move in with other people even for a little while? Y/N</li><li>■ Did you stay at a shelter, in an abandoned building, an automobile or any other place not meant for regular housing, for even one night? Y/N</li></ul> <p>It may be possible to determine whether current respondents were ever homeless and to link that information to NLS data that have already been collected.</p> <ul style="list-style-type: none"><li>■ Was there any time during your lifetime in which you:<ul style="list-style-type: none"><li>- Lived with others due to cost?</li><li>- Lived in places not intended for habitation?</li><li>- Lived in an emergency shelter?</li><li>- Lived on the streets (including car, campsite)?</li></ul></li></ul> <p>If yes to any of the above:</p> <ul style="list-style-type: none"><li>■ When did it occur? (month/year)</li><li>■ Who were you living with at the time:<ul style="list-style-type: none"><li>- Living alone</li><li>- Partner/spouse</li><li>- Children</li><li>- Other family member(s) (e.g., mother, cousin)</li></ul></li><li>■ How long did you live there?</li></ul>
--	---

\* These are modified versions of questions asked in the Fragile Families Study.