



Counting Couples, Counting Families



Family Structure and Instability Measures in Federal Data Collection Efforts

Kathleen Mullan Harris
University of North Carolina at Chapel Hill

This paper was presented at the *Counting Couples, Counting Families* 2011 Research Conference sponsored by the National Center for Family & Marriage Research, which is funded by a cooperative agreement, grant number 5 UOI AEOOOO01-04, between the Assistant Secretary for Planning and Evaluation (ASPE) in the U.S. Department of Health and Human Services (HHS) and Bowling Green State University. Any opinions and conclusions expressed herein are solely those of the author(s) and should not be construed as representing the opinions or policy of any agency of the Federal government.

Family Structure and Instability Measures in Federal Data Collection Efforts

Prepared for the Counting Couples Conference, July 19-20, 2011, NIH Campus

Sponsored by the National Center for Family & Marriage Research

Kathleen Mullan Harris

University of North Carolina at Chapel Hill

This document assesses the availability and quality of existing measures of family structure and family instability in federal data collection efforts in the U.S. A large body of research has documented the profound changes in contemporary family forms beginning in the 1960s when divorce rates began to rise and into the 1970s when nonmarital fertility rates increased, fueling the growth in single-parent families in the latter half of the 20th century. Soon after, research documenting the trends in cohabitation brought attention to the rise in cohabiting-parent families in the 1990s. Family scholarship focused primarily on the causes and consequences of family structure change to better understand the impact on children, as nearly universal were the findings that children living in two biological-parent families have better physical, emotional, and mental health as well as social and behavioral outcomes than children in all other non-traditional family structure forms (except, perhaps, single parent families formed through parental death). Underlying this general finding is the fundamental mechanism that living with both biological parents reflects greater stability in the family, neighborhood, school, peer, and socioeconomic context, whereas any non-traditional family form reflects some kind of family disruption and associated instability. These profound changes in the contexts in which children are raised and the salience of these family contexts for their lifecourse social, behavioral, and physical well-being make it a priority for federal data collections to measure family structure and family structure changes across time and among individuals in the U.S.

Almost all of current nationally representative studies that sample households or interview individuals include some measures of the family structures in which individuals in households live based on the data collection efforts in the crosswalk prepared by the National Center for Family and Marriage Research (NCFMR). The amount and quality of measures on family structure depend on several key criteria of the data collection effort: 1) study design; 2) scientific objective of the study; 3) data reporters in the family of interest; and 4) survey content.

The design of the study contains several elements. The overall design of the study can be either cross-sectional or longitudinal. By cross-sectional, the study does not follow the same cohort of individuals or families over time, but does interview repeated cross-sections of the U.S. population that represent different cohorts of individuals or families over time. Such cross-sectional studies can provide longitudinal data on family structure for individuals by collecting retrospective family structure histories, for example (more on that below). Longitudinal studies typically follow a cohort of individuals prospectively through time, where cohort membership is defined by experiencing the same event in a year or specific time period (e.g., birth, childbirth, grades in school) and are referred to as prospective longitudinal cohort studies. Longitudinal cohort designs have the advantage of capturing change in family structure as it occurs through time and therefore provide data with which to measure family instability as an individual experiences it and relate family instability to changes in outcomes. Cross-sectional studies included in the NCFMR crosswalk are the American Community Survey (ACS); Current Population Survey (CPS); National Survey of Family Growth (NSFG); and the Census. Longitudinal studies included in the NCFMR crosswalk are the Birth Cohort of the Early Childhood Longitudinal Study (ECLS-B); the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K); Fragile Families; Add Health; National Longitudinal Study of

Youth 79 (NLSY79); and the National Longitudinal Survey of Youth 97 (NSY97). Note that the Consumer Expenditure Survey (CE) and Survey of Income and Program Participation (SIPP) have an annual cross-sectional design with brief longitudinal panels within the year (CE) or for some cross-sections (SIPP).

The scientific objective of the study reflects what amount of survey time can be devoted to collecting information on family structure and family structure change. For example, the scientific goal of Fragile Families and the NSFG is to study the causes and consequences of family formation and family behavior; whereas Add Health is an omnibus study of the health status and health behavior of young people over their life course and the NLSY is a study focused primarily on the development of labor market skills and experiences over the life course. The data reporters about family structure are conditional on the design and purpose of the study as well. When children are young, as in ECL-B and ECLES-K or Fragile Families, parents report on family structure. When children are adolescents or young adults, they report on their family structures growing up as in Add Health or NLSY. Having two reporters of family structure is advantageous for validity checks and more complete measurement, such as mothers and fathers (small subset in Fragile Families) or parents and children (Add Health). Survey content is linked to the scientific objective of the study and is the main criteria by which questions on family structure vary displayed in the NCFMR crosswalk.

Measures of Interest

There are two concepts of family measures considered in this review: family structure and family instability. Both concepts may be measured by the same data, but often family instability requires additional data for more refined measurement involving change. I refer to *family structure status* as a static measure of family structure at a point in time (or point in child's life course). I refer to *family instability* as an indicator of change in family structure over time (or over developmental stages of a child's life, and ideally over the first 18 years of life when children live at home). Measures of family structure status and family instability can range from static, crude measures of resident parents' marital status to elaborate complex measures of the social and biological configuration of all family members inside and outside a household and changes in these configurations over time. Below I outline the various family structure measures for both status and instability from the ideal to less than ideal, and discuss the required sources of data for measures and their availability and quality. All measures of family structure are considered from the child's point of view, whether they are reported by adults on their childhood growing up or by adults as parents with children. That is, "family structure" as a measure is only relevant for families with children.

Family Structure Status Measures

Driven by research indicating that the type of parents and their configurations in families matter for family relationships and individual outcomes, family structure status measures require data on the number and type of parent(s) with whom the child lives.¹ By type, I refer to the relationship of each parent figure in the home to the child, including biological, step, adoptive, foster, or surrogate parents—both biologically related and non-biologically related surrogates. If there are two parents in the home, crossing these parent types results in the most detailed measure of family structure with two parents (5x5, or 25 family structure types), which can then

¹ Note that most concepts and measures of family structure status center on co-residential families in households. Extensions of this concept of family structure to include family members or siblings who live in other households (e.g., non-residential biological parents, non-residential half siblings) are possible and mentioned later.

be added to single parent types in each category as well (+5=30 family structure types). However, some combinations of parent types are rare or not meaningfully distinct from other types, such as step dad, foster mom and foster mom, surrogate dad. With this level of ideal information, however, a meaningful detailed measure of family structure status might include the following structure types: 1) two biological parents; 2) two adoptive parents; 3) two foster parents; 4) bio mom, step dad; 5) bio dad, step mom; 6) two bio-related surrogate parents (e.g., grandparents, aunts and uncles); 7) two non-bio surrogate parents (e.g., two step parents, neighbors); 8) single bio mom; 9) single bio dad; 10) single bio surrogate (e.g., sibling); and 11) single non-bio surrogate.

This level of detail would be especially informative for descriptive cohort and period trends in family structure, but may not be needed for specific research questions and analysis. For example, if the research focus is on variations in social and economic resources by family structure, the types of social parent figures may be less important than the number and biology of parents and collapsing family structure categories into five categories with three types of two-parents families (two bio, step, and surrogate) and 2 types of one-parent families (single bio parent, single non-bio) might be sufficient.

Research has also indicated that the marital status of two-parent families is salient for family relationships and outcomes for youth, and thus categorizing family structure types by marital vs cohabiting status to create subcategories of two bio, step, and surrogate parents may be useful. In addition, there is increasing interest in same-sex couples with children, another potential subcategory of two-parent families. Here I want to make an important point: sound and influential research involves a balance between fundamental detail and parsimony in measurement and modeling decisions. Measurement of such concepts as family structure depends on the specific research questions under study. Beginning with the most detail as possible, and then using theory and prior evidence to refine the measure to address specific questions is a good strategy. In sum, data that are needed to adequately define family structure status based on the current state of literature are number and sex of parent figures in the home, marital/cohabiting status of parent figures in the home, and the relationship of the child to each parent figure in the home.

Turning to the datasets in the NCFMR crosswalk, all of the cross-sectional data collection surveys, including the ACS, CE, CPS, NSFG, SIPP, and U.S. Census, have a main survey household respondent (HHR) list all members of the household on a roster, indicate each household member's current marital status (above age 14 or so) and what their relationship is to the HHR. Typically the HHR is a parent figure and thus these data allow the researcher to identify parent figures in the home (spouse or unmarried partner of HHR) but *only* the relationship between the HHR and any child listed on the household roster. This means that only one type of parent figure can be identified for two-parent families and the spouse or unmarried partner relationship with each child is unknown. So, for example, a male HHR who fills out the household roster may indicate that he is separated, has an unmarried partner in the household, and a biological child and adopted child. It is not clear whether the biological child to the HHR is also the biological child of the unmarried partner, whether the adopted child is the adopted child of the HHR, the unmarried partner, or both parent figures. Moreover, a married HHR who has a biological child in the household and lives with his wife could represent a two-biological parent family or a two-parent step family. Therefore, from the perspective of the child, complete family structure status cannot be determined and mainly crude measures of two- and one-parent family structure categories are possible. The biological relationship between all parents and children in a household cannot be determined, nor can complex blended family structures be specified.

In addition, relationship to HHR data varies across surveys. For example, in the ACS, a HHR's relationship to a child can be biological, adopted, step, or foster. The U.S. Census identifies biological, adopted, or step children of the HHR, and CPS only identifies foster or "child" of the HHR. If the HHR is not a parent figure in the household, family structure types are even more difficult to measure. Thus, if a grandparent in a three-generation household fills out the household information, parents will be listed in the household as sons or daughters to HHR, and their children as a grandchild of HHR, but the relationship between the parent(s) and children cannot be discerned.

The major limitation of these data are that the biological and social relationships between parents and children can only for determined for one parent in the household, thus limiting refinement in family structure types. This limitation is primarily due to the fact that relationship data are collected from the perspective of the household head who is filling out the survey, a design primarily driven by the purpose of the survey with greater emphasis on household composition, status, and behavior. To get complete information on the types of all parents in the household, the survey needs to collect relationship data from the perspective of the child.

The longitudinal studies do a better job of measuring family structure status that enables the researcher to sort out two-biological parent families from other types of two-parent families or blended families using different strategies. Fragile Families and Add Health gather a household roster and obtain the social and biological relationships between all parent figures and children in the household. In Fragile Families, the relationship between the child (i.e., nonmarital birth) with both the parent figures in the home are reported; and in Add Health, the adolescent respondent (and young adult in subsequent waves) reports on his or her relationship with every household member. Thus, complete information for the ideal family structure status measures is available in these two data sources. In addition, multiple reporters (fathers in Fragile Families for a subset, and parents in Add Health) allow for validity checks or to fill in missing data, making these data of very high quality. Marital and cohabiting status of parents, and mother's relationship to biological father are available from other questions in Fragile Families and from the parent interview in Add Health or can be determined by the type of parent partner listed (spouse, partner/boyfriend/girlfriend) in the household roster.

NLSY79 and NLSY97 also gather a household roster, but only obtain the relationship of HHR (NLSY79) or parent (NLSY97) to household members, similar to the census-based surveys. To get around this limitation, in the NLSY97 there are a series of follow-up questions that ask the parent respondent to identify the mother and father figure of the child who live in the household, the type of mother and father figure (bio, step, adopted, foster, or guardian), and whether the parent respondent has a spouse or partner in the household. This information enables the researcher to measure more refined family structure status as two bio parents, step families (1 bio, 1 step) or single bio-parent families, as well as single and two-surrogate parent families, getting close to the ideal measure described above.

ECLS-B and ECLS-K collect data on family structure by asking the parent respondent what their relationship is to the child (bio, step, adopted, foster, and other) along with other questions on the marital/cohabiting status of the parent respondent and whether the spouse or partner is the other biological parent. As with NLSY97, one and two parent families can be delineated according to biology in the ECLS surveys.

In sum, the most detailed information with which to measure family structure status comes from data on the social and biological relationships between parents and children and the marital/cohabiting status of parent figures in the household. Ideally, this information can be

obtained from the child's reports of relationships with members in a household roster, such as in Add Health, but can also be obtained from one parent reporter in Fragile Families.² Alternatively, parent reports of their relationship with children provide half the structure, and along with data on the presence of spouse or partner in a household, and on the biological relationship of spouse or partner to the child, close to the ideal measure of family structure status is possible. This form of data is available in NLSY97, ECLS-B, and ECLS-K, but is harder to work with from a researcher's perspective, does not allow as much detail, and for these reasons of somewhat lower quality. Family structure measures from the cross-sectional and census-based surveys are of lower quality because the social and biological relationships between both parents in two-parent families are unknown, leading to less detailed measures of family structure status along dimensions that previous research has indicated are critical to outcomes. Exceptions include SIPP and NSFG, which include some supplemental questions on children's living arrangements, and the relationship of the HHR's spouse or unmarried partner to the child.

Family Instability Measures

In the most general sense, family instability is measured by family disruption or transitions and change in family structure over time, where time refers to the child's life course. A crude indicator of family instability is parents' marital status, where divorced, separated, or widowed indicates some prior disruption. But any disruption is not a satisfying measure given the research indicating that the number of transitions may matter, the life stage in which disruptions occur may matter, and the length of time since disruption may matter for children's adjustment and well-being. Thus, there are several concepts related to family instability that require more refined measures and detailed data.

There are two dimensions relevant here: 1) change—when, what, and how often; and 2) duration—length of time since last family structure change and length of childhood life course in different family structure types. The number of family structure changes is an overall measure of family instability, and can either be measured across the life course, within different life stages (i.e., childhood, middle and late adolescence) or up to the date of survey collection. The type of change may be relevant for the theoretical research question under study. That is, change from a two-parent family to a one-parent family may be detrimental whereas change from a one-parent to a two-parent family beneficial for children and parents; or change from a two bio-parent family may be more detrimental than change from a two-parent step family. Lastly, family structure change, as well as the number and type of changes, may have differential impacts on children and parents depending on the developmental life stage of the child in which change occurs. The duration dimension captures the stability of family structure experiences, and two measures are common. Because some evidence suggests that children's adjustment to family structure change improves with longer time in the new family structure, the duration since family structure change is a salient measure. To test theoretical hypotheses about the impacts of different family structure statuses, the length or proportion of the child's life spent in each family structure type is a high quality but data demanding measure.

All of the instability measures require longitudinal data to capture change over time and can be obtained using two survey approaches. One approach is to gather retrospective data on family disruptions, either with a set of questions (i.e. how many times have you been married or living with a partner; number and duration parents' previous marriage—NLSY97) or with a

² It is not clear from the crosswalk whether one parent in Fragile Families reports on relationships between the child and both mother and father figures in the home, or both parents report on the relationships. The crosswalk does indicate that the relationship between both the mother and the father with the child is known, however.

marriage and cohabitation history of the parent with whom the child lives. Both cross-sectional and longitudinal study designs can provide these types of data. A second approach is to collect a household roster indicating the relationships between children and household members in waves of data collection on a prospective cohort and measure change in parent figures over time. Only the longitudinal studies can provide this type of information.

Turning to the cross-sectional surveys in the NCFMR crosswalk, it does not appear that any of these datasets provide a marriage or cohabitation history of a parent because these studies are primarily focused on current statuses.³ So, the only available measures of family instability for the census-based cross-sectional studies (ACS, CE, CPS, SIPP, U.S. Census) come from current marital and cohabitation status where any disruption can be crudely measured. One exception is SIPP, which includes some additional questions for why new household members “joined” the household, with possible responses of marriage, separation or divorce or birth. This information allows for some dynamic measure of family structure change, but certainly not longitudinal.

In the longitudinal studies, repeated measures of relationships between parents and children based on the household roster and marital/cohabiting status of the HHR across the waves of data collection make it possible to identify change in types of parent figures over time and construct measures of family instability. Change in family structure types will be of higher quality if the periodicity of the survey is often, thereby capturing changes that may be missed with longer intervals between survey waves of more than 3 years. Note that data quality on instability measured with this approach is dependent on the data quality of the family structure status measure reviewed above, because it is this repeated measure that is the basis of the instability measure. Thus, the highest quality data on instability using the household relationships approach will be found in Fragile Families, ECLS-B, ECLS-K, and NLSY97 because they collect data from the same cohort every 1-3 years (last wave of Fragile Families had a 4-year interval).

Note that creating these instability measures are time-intensive and complex because the household relationship data are only available for one parent (except in Fragile Families), and must be supplemented by additional questions with complex skip patterns to create the entire array of social and biological relationships with children in two-parent families, and then these are compared across time. Another limitation to this approach is that without the parent’s marital or cohabitation history, the child’s experience of family structure change may be left-censored by the date at which the survey begins in relation to the age of the child. This is not a problem for ECLS-B and Fragile Families because these surveys began with the birth of the index child. But with NLSY97 and ECLS-K, the family structure history of the child is only known from the age of the child at the first interview wave. Thus, change measures of number of family structure changes, types of changes, and life stages in which changes occur may be incomplete or underestimated. Similarly, duration measures will also be truncated if the entire child’s life, at least up to the current survey point, cannot be included in the measurement.

The other approach of using the marital and cohabitation history of the mother or the custodial parent, supplemented with data on the relationship between the parent’s partners and the child, or at least which partner is the biological parent of the child, can overcome these limitations. Based on my perusal of these study websites, it appears that only Add Health has

³ The NSFG website indicates there is a marital and cohabitation history, but it was not indicated on the crosswalk. The marital and cohabitation history makes it possible to create instability measures in NSFG, but only for the residential parent respondent, thus changes in parent figure types over time is not known.

both a marriage and cohabitation history of the resident parent.⁴ Knowing Add Health best, the parent's marriage and cohabitation history covered the adolescent respondent's life from birth to the time of the first interview for more than 90% of the adolescent respondents. With data from the household roster on the social and biological relationships to every household member, on when the adolescent ever lived with the non-residential biological parent if not living with him or her at the first interview, all instability measures are available, including the number and type of transitions, the timing of transitions, and the duration of the life course that the child lived in different family structures. ECLS-K does not have a marital or cohabitation history of the parent and NLSY97 only asks the youth respondent the number and duration of the parents' previous marriages. Note, though, that there are always data issues with complex change measures such as these, including missing data, inconsistencies, and highly variable and unstable living arrangements of children (such that children are not always living within the parent's measured family structures across time), and these problems are greater for children and parents with the most instability. However, the datasets with the richest measures of family structure status (Add Health, Fragile Families, NLSY97, and ECLS), with short periodicity (Fragile Families, ECLS, NLSY97) or high quality retrospective data on parents' marital and cohabitation histories (Add Health), provide immense opportunities for research on the dynamics of family structure and family instability and their consequences for children and parents.

Other Family Structure Measures

These federally-funded national data collection projects provide other types of data to allow extensions of family structure measures and I will mention a few. All of the longitudinal studies and NSFG and SIPP provide information on non-residential biological parents, their living arrangements and some history of when they ever lived with their biological child. With these data more nuanced measures of family structure that represent the complete configuration of social and biological parents with whom the child has access can be created. So, for example, a child may live with a bio mom and step dad, but also be in contact with or sometimes live with his or her non-residential bio dad and step mom and therefore have access to 4 parent figures compared to a child who lives with a single parent and has no contact with the non-resident biological parent. Fragile Families, NSFG, and SIPP, in particular, collect additional information on the living arrangements of children, especially if children live in multiple households during the year.

Co-residential intergenerational family structures can also be identified for all surveys that collect a household roster (all data collections in the crosswalk; ECLS collects a household roster in the first interview, then updates it in subsequent waves). The presence of grandparent(s) in the household represent greater access to biologically-related adults. Some surveys ask particularly about grandparent relationships (ACS, CE).

Most of the longitudinal studies in the crosswalk provide information on biological children who are not living with the parent, which is important for calculation of total number of children and number of siblings. Fragile Families and the NLSY surveys provide the most information on non-residential or other biological children of parents. Add Health provides information about the biological relationships among siblings, and this is because the relationship information based on the household roster is gathered from the perspective of the child rather than the parent. Siblings can be classified as twin full, non-twin full, half, and step sibs, and identical twins are differentiated from fraternal twins. Other surveys allow the

⁴ NLSY79 collects a marital history of the youth (14-21 in 1979) as he/she ages, which is useful for studying family structure change among the children of the NLSY79.

identification of siblings in the household, though their biological relationship is not always known. These data provide even more refined measures of family structure according to the sibship relations, also found to vary with outcomes for children.

Finally, many of these surveys have contextual information measured at multiple spatial units that are attached to individual records. There are various measures of family structure and instability that can be defined at the neighborhood, school, and peer levels. Many studies have attached census measures on the proportion of married couples with children and the proportion of females headed families with children measured at the block group or census tract, for example. Add Health has particularly unique contextual data that measures the family structure of families in the school and the peer groups in which adolescents are embedded, in addition to the neighborhood measures. Thus, the proportion of the adolescent's friends who live in a single-parent family (based on actual peer reports of their family structure) and the proportion of students in the school who live in single-parent families represent family structure norms and models in the social environments in which children live.