

This article was downloaded by: [University of Minnesota Libraries, Twin Cities]

On: 06 November 2014, At: 09:54

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Historical Methods: A Journal of Quantitative and Interdisciplinary History

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/vhim20>

### IPUMS-CPS: An Integrated Version of the March Current Population Survey, 1962-2002

Miriam L. King<sup>a</sup> & Michele Tertilt<sup>a</sup>

<sup>a</sup> Minnesota Population Center, University of Minnesota

Published online: 30 Mar 2010.

To cite this article: Miriam L. King & Michele Tertilt (2003) IPUMS-CPS: An Integrated Version of the March Current Population Survey, 1962-2002, *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 36:1, 35-40, DOI: [10.1080/01615440309601213](https://doi.org/10.1080/01615440309601213)

To link to this article: <http://dx.doi.org/10.1080/01615440309601213>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

# IPUMS-CPS

## An Integrated Version of the March Current Population Survey, 1962–2002

MIRIAM L. KING

MICHELE TERTILT

*Minnesota Population Center*

*University of Minnesota*

**Abstract.** The Minnesota Population Center (MPC) is creating an integrated version of the March U.S. Current Population Survey (CPS) covering the years 1962–2002. The CPS is one of the few comprehensive data sources available for researching social and economic trends between decennial censuses. The survey also includes useful information—such as employment patterns and participation in social welfare programs—that is not available in the census. The new database will be compatible with the decennial census data available in the Integrated Public Use Microdata Series (IPUMS). For variables found in both sources, the integrated CPS will use IPUMS codes, and the documentation and data access system for the CPS data will be based on IPUMS principles and infrastructure.

**Keywords:** Current Population Survey (CPS), demography, economics, IPUMS, microdata survey

This article describes the effort of the Minnesota Population Center (MPC) to develop a harmonized version of the U.S. Current Population Survey (CPS) for the period from 1962 to 2002. The CPS is a powerful source for investigating social and economic trends between decennial censuses. The new data series—IPUMS-CPS—will be compatible with the Integrated Public Use Microdata Series of U.S. decennial census data (IPUMS-USA).

Whereas researchers and policy analysts have made extensive use of the CPS, relatively few scholars have used the full range of survey years to study long-term economic and social change. Fewer still have taken advantage of the considerable overlap in content between the public use samples of the U.S. census and the CPS. Because many of the same questions are included in the decennial census and the CPS, the latter source is ideal for providing information on years between the census benchmarks. IPUMS-CPS is designed to facilitate such research by providing harmonized coding schemes, consistent data formatting, thorough documentation, and free easy-to-use Web access that has been the hallmark of previous IPUMS projects.

### Background

The CPS is a monthly U.S. household survey currently covering about 56,000 occupied housing units and conducted jointly by the U.S. Census Bureau and the Bureau of Labor Statistics (BLS). Initiated in the 1940s in the wake of the Great Depression, the survey was designed to measure unemployment. Providing monthly statistics on unemployment is still one of its primary purposes. A battery of labor-force and demographic questions, known as the “basic monthly survey,” is asked every month. Over time, the Census Bureau and the BLS have added supplemental monthly inquiries on special topics, including health, education, work experience, childcare, veterans’ issues, computer use, and marriage and fertility. Among these supplemental surveys, the March Annual Demographic File and Income Supplement (hereinafter referred to as the March CPS) is most widely used by social scientists and policymakers.

Users of the public use files of March CPS data who study long-term change must cope with repeated shifts in coding schemes, record layouts, universe definitions, and variable availability, as well as a complicated hierarchical record structure. Documentation is sparse, and discussions of comparability issues are limited. Particularly scanty is the documentation for the survey data from 1962 to 1967, which was preserved by researchers at the University of Wisconsin rather than at the Census Bureau. The data collected before 1962 were not preserved at all.

With funding from the National Science Foundation, the MPC is creating a harmonized version of the March CPS from 1962 to the present (IPUMS-CPS). This data set is compatible with IPUMS data from the U.S. decennial censuses of 1960, 1970, 1980, 1990, and 2000. Researchers can thus take advantage of the relatively large sample sizes of the IPUMS at 10-year intervals and fill in information for the intervening years using IPUMS-CPS. Our first data release, scheduled for 2003, will consist of variables that are simultaneously available in the IPUMS and the March

CPS. Table 1 presents these variables, together with the years that they are available for the March CPS. A second data release will cover those March CPS variables that have no equivalent in the decennial census data. Data and documentation for IPUMS-CPS will be freely available through the IPUMS data access system at <http://www.ipums.org>.

This attempt is not the first to simplify use of these data. In the late 1980s, Robert Mare and Christopher Winship created the widely used Mare-Winship Uniform March Files, which include recodes of 168 variables from the March CPS of 1962 to 1988. These files were later extended to 1992 by Christine Collins and are now distributed by the National Bureau of Economic Research ([http://www.nber.org/data/mare\\_winship.html](http://www.nber.org/data/mare_winship.html)).

More recently, the Unicon Research Corporation (2001) developed CPS Utilities, a software system that provides access to all surviving CPS data. Part of the system provides tools to dynamically recode approximately 80 of the March CPS variables into a least common denominator that allows comparison over time.

Although the Mare-Winship files and CPS Utilities represent valuable contributions, their approach differs substantially from that of IPUMS-CPS. The Mare-Winship files lost much of the detail in the original CPS files and did not provide any documentation of comparability issues. The massive Unicon project covers far more data than IPUMS-CPS, but it is essentially a tool for accessing the CPS in its original form and makes only modest efforts at harmonization over time. IPUMS-CPS is much more interventionist: the data are completely reformatted and recoded, redundant items are eliminated, and errors in the original data are corrected whenever possible. Moreover, unlike the previous projects, IPUMS-CPS is concerned not only with the compatibility of CPS files across time but also with harmonization of the CPS and the decennial census.

### Differences between the March CPS and IPUMS

Some differences between the March CPS and the IPUMS census data are inherent in the distinction between surveys and censuses. Although the U.S. census is conducted only once per decade, the March CPS is available for every year beginning in 1962. Whereas the census seeks universal coverage, the CPS draws on a relatively small number of households. Sample size for the March CPS varies over time, ranging from about 50,000 to roughly 180,000 persons. Designed to measure the labor-force participation of the civilian population, the CPS excludes some groups covered by the census, such as inmates of institutions and U.S. citizens living abroad. Members of the armed forces, except those living in family housing, are also excluded. Before 1968, children under 14 were not included in the survey, although summary information about the number and approximate ages of children was provided.

The sample designs for the CPS differ from those used for the decennial census. The CPS is a rotating panel survey: housing units are interviewed monthly for four consecutive months, leave the sample for eight months, and then are interviewed four more times.<sup>1</sup> The fact that there is a 50 percent overlap in the addresses surveyed across 16 months allows some analysis of particular households over time. Some questions in the survey are posed only to those units that have been in the sample for a given number of months.

In all years, the CPS used a multistage stratified sampling scheme to draw representative samples in each state and the District of Columbia. The United States is divided into primary sampling units (PSUs) consisting of metropolitan areas or groups of counties; PSUs within states are grouped into homogeneous strata by their economic and social characteristics; one PSU within each strata for each state is selected; and approximately four households (defined by the address of the housing unit) are then selected for interviewing. In some years, the CPS includes oversamples of population subgroups of special interest. In addition, the sampling weights are adjusted to accommodate different probabilities of nonresponse for various demographic groups and to bring the sample population in line with the age, sex, and race distribution of the entire population.

Because of the complexity of the sample design for the CPS, the sample weights are more complicated than those for census data. There are several sets of weights from which to choose. Their proper selection depends on which questions are the subject of inquiry and whether the units of analysis are households, families, or individuals. Our documentation will include full discussion of the use and comparability of all sample weights.

### Harmonized Coding Schemes

To create harmonized codes for CPS variables, we followed the same procedures used in the IPUMS project for U.S. census data and the IPUMS-International project (see Esteve and Sobek in Part Two of this issue). The harmonization was designed to meet two goals. First, we wanted users to be able to move smoothly between the IPUMS-USA samples when the same variable was included in the March CPS. In these cases, we recoded the March CPS data directly into the IPUMS-USA coding scheme. Second, we wanted to create harmonized codes for variables unique to the March CPS; these cases required new coding schemes.

To harmonize the CPS, we developed data transformation tables for each variable similar to those developed for IPUMS-USA (Hall et al., 1999) and IPUMS-International (see Esteve and Sobek in Part Two of this issue). Each table provides information on the location of the original variable in each sample, original codes, and the new IPUMS-CPS codes. Unlike the census, however, the values and locations for a given variable in the CPS are often the same for a series of years. For this reason, we modified the translation-

TABLE 1. Harmonized Variables in CPS and IPUMS, by Year (1960–2002)

Variable name	Meaning	Year	
		CPS (1962–2002)	IPUMS (1960–2000)
ABSENT	Absent from work last week	1962–2002	1980, 1990, 2000
AGE	Age	1962–2002	1960, 1970, 1980, 1990, 2000
AGEMARR	Age at first marriage	1962, 1965, 1967–71	1960, 1970, 1980
ANCESTR	Ancestry	1971–1975	1980, 1990, 2000
AVAILBLE	Available for work	1968–1993	1980, 1990, 2000
BIRTHQTR	Quarter of birth	1968–1971	1960, 1970, 1980
BPL	Birthplace	1994–2002	1960, 1970, 1980, 1990, 2000
CITIZEN	Citizenship status	1994–2002	1970, 1980, 1990, 2000
CLASSWKR	Class of worker	1962–2002	1960, 1970, 1980, 1990, 2000
COLL5YR	College 5 years ago	1980	1970, 1980
DISABWRK	Work disability	1988–2002	1970, 1980, 1990, 2000
EDUC99	Educational attainment, 1990	1992–2002	1990, 2000
EDUREC	Educational attainment recode	1962, 1964–2002	1960, 1970, 1980, 1990, 2000
ELDCH	Age of eldest own child in household	1968–2002	1960, 1970, 1980, 1990, 2000
EMPSTAT	Employment status	1962–2002	1960, 1970, 1980, 1990, 2000
FAMSIZE	Number of own family members in household	1968–2002	1960, 1970, 1980, 1990, 2000
FAMUNIT	Family unit membership	1962–2002	1960, 1970, 1980, 1990, 2000
FARM	Farm status	1962, 1977–1992	1960, 1970, 1980, 1990, 2000
FBPL	Father's birthplace	1994–2002	1960, 1970
FOODSTMP	Food stamps reciprocity	1982–2002	2000
HIGRADE	Highest grade of school	1962, 1964–1991	1960, 1970, 1980
HISPAN	Hispanic origin	1968–2002	1970, 1980, 1990, 2000
HRSWORK1	Hours worked last week	1962–2002	1980, 1990
HRSWORK2	Hours worked last week, intervalled	1962–2002	1960, 1970, 1980, 1990
INCBUS	Nonfarm business income	1962–2002	1970, 1980, 1990, 2000
INCFARM	Farm income	1962–2002	1970, 1980, 1990
INCINVT	Interest, dividend, rental income	1968–1975	1980, 1990, 2000
INCRETIR	Retirement income	1976–2002	1990, 2000
INCSS	Social Security income	1968–2002	1970, 1980, 1990, 2000
INCSUPP	Supplemental Security income	1976–2002	2000
INCTOT	Total personal income	1962–2002	1960, 1970, 1980, 1990, 2000
INCWAGE	Wage and salary income	1962–2002	1960, 1970, 1980, 1990, 2000
INCWELFR	Welfare (public assistance) income	1968–2002	1970, 1980, 1990, 2000
IND	Industry	1962–2002	1960, 1970, 1980, 1990, 2000
IND50	Industry, 1950 basis	1968–2002	1960, 1970, 1980, 1990, 2000
LABFORCE	Labor-force status	1962–2002	1960, 1970, 1980, 1990, 2000
LOOKING	Looking for work	1962–2002	1980, 1990, 2000
MARRQTR	Quarter of first marriage	1968–1971	1960, 1970, 1980
MARST	Marital status	1962–2002	1960, 1970, 1980, 1990, 2000
MBPL	Mother's birthplace	1994–2002	1960, 1970
METAREA	Metropolitan area	1962–2002	1970, 1980, 1990, 2000
METRO	Metropolitan central city status	1963–2002	1960, 1970, 1980, 1990, 2000
MIGCITY5	SMA central city 5 years ago	1980, 1985, 1995	1980, 1990
MIGTYPE5	Metropolitan status 5 years ago	1980, 1985, 1995	1960, 1970, 1980, 1990
MIGMET5	Metropolitan area of residence 5 years ago	1980, 1985, 1995	1980, 1990
MIGPLAC5	State or county of residence 5 years ago	1980, 1985, 1995	1970, 1980, 1990, 2000
MIGRATE5	Migration status 5 years ago	1980, 1985, 1995	1960, 1970, 1980, 1990, 2000
MILIT5YR	Armed forces 5 years ago	1980	1970, 1980
MOMLOC	Mother's location in household	1968–2002	1960, 1970, 1980, 1990, 2000
NATIVITY	Generation: nativity recode	1994–2002	1960, 1970
NCHILD	Number of own children in household	1968–2002	1960, 1970, 1980, 1990, 2000
NCHLT5	Number of own children under 5 in household	1968–2002	1960, 1970, 1980, 1990, 2000
OCC	Occupation	1962–2002	1960, 1970, 1980, 1990, 2000
OCC1950	Occupation, 1950 basis	1968–2002	1960, 1970, 1980, 1990, 2000
OWNERSHP	Ownership of dwelling	1976–2002	1960, 1970, 1980, 1990, 2000
PHONE	Telephone availability	1983–2002	1960, 1970, 1980, 1990, 2000

(table continues)

TABLE 1. (Continued)

Variable name	Meaning	Year	
		CPS (1962–2002)	IPUMS (1960–2000)
POPLOC	Father's location in household	1968–2002	1960, 1970, 1980, 1990, 2000
POVERTY	Poverty status	1963–2002	1960, 1970, 1980, 1990, 2000
RACE	Race	1962–2002	1960, 1970, 1980, 1990, 2000
REGION	Census region and division	1962–2002	1960, 1970, 1980, 1990, 2000
RELATE	Relationship to household head	1962–2002	1960, 1970, 1980, 1990, 2000
SEX	Sex	1962–2002	1960, 1970, 1980, 1990, 2000
SPLOC	Spouse's location in household	1962–2002	1960, 1970, 1980, 1990, 2000
STATEFIP	State (FIPS code)	1962–2002	1960, 1970, 1980, 1990, 2000
UHRSWORK	Usual hours worked per week	1976–2002	1980, 1990, 2000
UNITSSTR	Units in structure	1979–2002	1960, 1970, 1980, 1990, 2000
WKSUNEMP	Weeks unemployed last year	1962–2002	1980
WKSWORK1	Weeks worked last year	1976–2002	1980, 1990, 2000
WKSWORK2	Weeks worked last year, intervalled	1962–2002	1960, 1970, 1980, 1990, 2000
WORK5YR	Working 5 years ago	1980	1970, 1980
WORKEDYR	Worked last year	1962–2002	1960, 1970, 1980, 1990, 2000
YNGCH	Age of youngest child in household	1968–2002	1960, 1970, 1980, 1990, 2000
YRIMMIG	Year of immigration	1994–2002	1970, 1980, 1990, 2000

Note. SMA = standard metropolitan area. FIPS = Federal Information Processing Standards.

table format from that used in other IPUMS projects. Instead of devoting a separate column to each data file, we created a new column only when the location or the coding for the variable changed. Each column is used for a range of CPS files and specifies the beginning and ending year.

An example may make this clearer. Consider the variable POPSTAT, which indicates whether the individual is an adult civilian, a member of the armed forces, or a child. These subgroups are consistently coded as 1, 2, and 3, respectively. In the public use files of the March CPS, this variable occupies column 102 of the person record from 1976 to 1987 and column 26 of the person record thereafter. The definition of *child* also changes from under 14 years to less than 15 years in 1980. The translation table for POPSTAT in this period has three columns: (1) 1976–1979 (location 102, child is under 14), (2) 1980–1987 (location 102, child is under 15), and (3) 1988–2002 (location 26, child is under 15). For more complicated cases (e.g., when no original code was supplied for cases that are “not in universe” or when data are drawn from more than one nonadjacent variable), we supplemented the translation tables with additional programming instructions.

Because the Census Bureau conducts both the decennial census and the CPS, there is some overlap in the subject matter, wording, and coding of questions from these two sources. For example, both the long-form census questionnaire and the CPS collect information on the current occupation of respondents. Since 1968, the CPS has used the occupational coding schemes of the census (e.g., elementary school teachers are given the occupational code of 182

in both the 1960 census and the 1968 March CPS). To accommodate researchers interested in long-term change, we have also recoded CPS occupational data (beginning with 1968) into the 1950 occupational classification system used by the IPUMS as a standard classification for all U.S. census years. From 1962 to 1967, the March CPS used simpler occupational classification schemes consisting of only about 35 categories. To allow consistent comparisons from 1962 to the present, we developed a simplified occupational coding scheme.

The ease of recoding CPS data into IPUMS coding schemes depends on the variable in question. Comparisons between IPUMS-USA and IPUMS-CPS fall into five categories. In the first case, the CPS data for a given variable have the same substantive values as a similar variable in the IPUMS, but the arbitrary numerical codes that represent those values differ. For example, the variable for citizenship status codes foreign-born persons who are not naturalized citizens as “3” in the IPUMS coding scheme and “5” in the March CPS data for the period from 1994 to 2002. In such cases, it is straightforward to recode the CPS values into the IPUMS coding scheme. In the second case, less-detailed codes are available in the CPS than in the census. There are, for example, 13 general codes for relationship to household head in the IPUMS, but for several years the corresponding variable in the CPS has had only five values. In such cases, using a subset of the values in the IPUMS variable, we recoded the CPS data. In the third case, more-detailed codes for a given variable are provided by the March CPS than in the census; additional codes are then added to the IPUMS coding

scheme to preserve such detail. For instance, telephone availability is a dichotomous variable in IPUMS-USA; the March CPS distinguishes between households that do not have a telephone available for use, those with access to a phone outside the household, and those with a phone inside the household. In the fourth case, variables are only superficially similar and require the creation of new IPUMS variables: for example, census respondents who served in the armed forces report all specific time periods of their service (such as World War II and the Korean War). In the CPS, veterans report only their most recent period of service. Here we have created new veteran status variables for decennial census data from 1960 to 2000 corresponding to those in the CPS.

A fifth situation, when variables are available exclusively in the CPS, is of particular benefit to researchers. The March CPS includes a large number of variables that have no equivalent in the decennial census or in the American Community Survey (a 3 percent sample of U.S. households that is scheduled to take the place of the census long form in 2003). Excluding data quality flags, the 1990 census data in IPUMS consist of 72 household-level variables and 111 person-level variables. The original public use version of the 1990 March CPS includes 108 household-level variables, 68 family-level variables, and 248 person-level variables (excluding data quality flags). The discrepancy is even greater in the most recent years because additional questions were added to the March CPS to investigate such topics as immigration and welfare-reform programs. Some of these CPS variables are merely recodes of information available elsewhere in the survey (such as age groups recoded from single-year age data). Other variables contain new information of considerable interest to investigators. The March CPS covers some topics unavailable in the U.S. census (see appendix). Particularly noteworthy are the CPS labor-force variables that report an individual's primary job and other labor-market experiences during the same year for which his or her income was reported. This information is especially important to economists because the decennial census reports on work from the previous week and income from the previous calendar year.

For March CPS variables that have no equivalent in IPUMS, we are using the same principles as IPUMS-USA and IPUMS-International to develop harmonized coding schemes. To maximize temporal compatibility of variables with no loss of detail, IPUMS-CPS employs a composite coding system. The first digits of the composite code provide information available across all survey years. One or two additional digits provide added detail for a particular survey year or group of years. Throughout the survey period, the March CPS has collected information about the reasons why employed persons were not at their job during the previous week. In the 1962–67 data sets, only four specific responses (illness, vacation, bad weather, and labor dispute) and a catchall “other” category were coded; by 2002, the data set included 11 specific responses. For the harmonized ver-

sion of this variable, we adopt a two-digit code. The first digit represents the codes present in the early data sets; the second digit provides the additional detail present in later years by elaborating on the “other” response. In contrast, other efforts to harmonize CPS data (e.g., by Mare and Winship at the University of Wisconsin and by the Unicon Corporation) have adopted a least-common-denominator approach when recoding variables to achieve consistency over time.

As noted above, many variables included in the CPS public use files are recodes of information recorded elsewhere in the survey. In general, these recodes are not included in IPUMS-CPS. For example, the March CPS family record for 1963–67 includes 92 variables. Twenty-five of these variables summarize characteristics of the head of the primary family (or head of household), and 14 summarize characteristics of the wife. Such summary variables, based only on recoding existing variables, add no information. The recoded variables were intended as a convenience for data analysts, but when there are multiple versions of the same variable, the data and documentation are actually more difficult to use. Because our coding system loses no information, researchers can recode variables to suit their particular analyses.

### Data Reformatting and Errors

Users of IPUMS or other census microdata are familiar with the hierarchical format of household records followed by individual records for each person. The original form of the CPS data sets follows a more complex hierarchical format. For most years (including the most recent), the CPS data sets have a three-part hierarchy: (1) a household record, followed by (2) family records, followed by (3) person records for each member of a family. Family records are supplied for each of the following: primary individuals (persons who head households but do not live with relatives); secondary individuals (persons who do not head households and do not live with relatives); secondary families (groups of related persons who are not related to the household head); and related subfamilies (married couples or parents and children related to the household head). For the period from 1963 to 1975, the original CPS data consist of family and person records only; for 1962, there are only person records.

IPUMS-CPS reformats the files for all years into a standard two-level hierarchy of household and person records. When there was no household record in the 1962 file, we created one using information from the person record of the household head. Data from family records were appended to the person record of each person who was a member of a given family. We removed the family record itself from the data.

Relatively little original data are included on family records; most variables are recodes from the individual records. We have preserved the unique family-record data

so that researchers can use them if they wish (e.g., to duplicate published research that used them), but we question their general accuracy. Several scholars have noted that the CPS family records severely underestimate the prevalence of related subfamilies consisting of a married couple or parent and children related to the household head (Graham and Beller 1985; Bianchi 1995; London 1998; Ruggles and Brower 2003). Our analysis of CPS variables summarizing family structure show additional problems. We found many discrepancies between the CPS family interrelationship variables and the relationship to head variable on the person record, such as 11-year-old children said to be the offspring of 12-year-old parents.

Drawing on our experience with the IPUMS project, we have constructed our own indicators of family interrelationships for IPUMS-CPS (see Ruggles 1995 for a description). These include family interrelationship pointer variables that allow researchers to link husbands to wives and parents to children. We are also constructing summary measures of family structure familiar to users of IPUMS census data, including family membership, family size, number of own children, number of own children under 5, and number of married couples in the household.

We have corrected and flagged a variety of other errors in the CPS data. Such mistakes are particularly common in data from the 1960s. Many survey questions were asked of only a subset of respondents, and "not in universe" codes were not always applied to those who were not questioned. Codes for individuals were occasionally inconsistent (e.g., persons younger than age 10 listed as members of the armed forces). For some years in which each household was supposed to have a unique identification number, the same number was assigned to more than one household. Our analysis revealed that some household members had been separated in the file, so we remerged these households.

### Documentation and Data Release

Perhaps the most valuable feature of IPUMS-USA is its comprehensive documentation. IPUMS-CPS follows this model. For each variable, the Web-based documentation will outline the general subject matter, the question wording, the universe of respondents, and the years for which the variable is available. To assist analysts who combine the decennial census data from IPUMS-USA with the annual data from IPUMS-CPS, the documentation will discuss any subtle differences between the CPS and the decennial census data. Comparability across CPS survey years will be covered, as will errors and inconsistencies in the original data and our efforts to correct them. Additional documentation will provide information on such topics as sample design, the use of weights, and the calculation of standard errors.

The IPUMS-CPS data will be made available in 2003. We anticipate that the availability of a harmonized version of CPS, covering 40 years of economic and demographic change, will spur new research efforts. In particular, the use of comparable coding schemes for the decennial census data in IPUMS-USA and for the annual survey data in IPUMS-CPS will encourage researchers to combine these important sources of social and economic data.

#### APPENDIX. Selected Topics Covered by March CPS Unavailable in the Census

Health insurance coverage  
 Participation in social welfare programs, including  
   Food stamps  
   Free school lunches  
   Energy assistance  
   Housing subsidy  
   Welfare reform programs  
 Hourly wages  
 Additional income categories  
 Tax liability  
 Union membership  
 Reason for, duration, and spells of unemployment  
 Actions taken to find work  
 Reason for not looking for work  
 Reason for moving/migration  
 Number of employees of employer  
 Work history during preceding year  
 Occupation, industry, and class of worker during preceding year

#### NOTE

1. Under the CPS sampling scheme, researchers select housing addresses and interview the residents. If those who live at a given address move out during the survey period and new residents move in, those new residents become part of the sample. Those persons who moved away drop out of the sample.

#### REFERENCES

- Bianchi, S. M. 1995. The changing demographic and socioeconomic characteristics of single parent families. *Marriage and Family Review* 20: 71–97.
- Graham, J. W., and A. H. Beller. 1985. A note on the number and living arrangements of women with children under 21 from an absent father: Revised estimates from the April 1979 and 1982 Current Population Surveys. *Journal of Economic and Social Measurement* 13: 209–14.
- Hall, P. K. et al. 1999. IPUMS metadata: Documenting 150 years of census microdata. *Historical Methods* 32: 111–18.
- London, R.A. 1998. Trends in single mothers' living arrangements from 1970 to 1995: Correcting the Current Population Survey. *Demography* 35: 125–31.
- Ruggles, S. 1995. Family interrelationships. *Historical Methods* 28: 52–58.
- Ruggles, S., and S. Brower. 2003. The measurement of household and family composition in the United States, 1850–1999. *Population and Development Review* 29(1): In press.
- Unicon Research Corporation. 2001. *The CPS utilities user's manual*. Santa Monica, Calif.: Unicon Research Corporation.