

# APPENDIX F

## User Notes for the SPD 2000 Minimally Edited Cross-sectional File

This section contains information relevant to the Survey of Program Dynamics 2000 minimally edited Cross-sectional File. For an updated list of user notes always refer to the U.S. Census Bureau's SPD Internet site at <http://www.census.gov/spd>. The user notes are part of the technical documentation which is found under "SPD Publications". In addition, some user notes are found under "SPD News User Notes". The Internet site will be updated as additional user notes become available.

### User Notes for the SPD 2000 Minimally Edited Cross-sectional File

#### *1. General Notes*

The data on the SPD 2000 Cross-sectional public use file is only minimally edited. The demographic variables are longitudinally edited and are indicated with an 'E' or an 'R' in the second right character of the variable name. Additionally, the income amount variables are edited only to the extent that multiple items are used as input to create the recoded amount on the file. All other variables are unedited and are indicated with an "U" in the second right character of the variable name. Variables that are topcoded for confidentiality reasons are indicated with an "T" in the second right character of the variable name.

Blank or missing values for numeric variables are 9-filled (9, 99, 999, 9999, 99999, 999999, 9999999) in the 2000 Cross-sectional Data Files. Consequently, during any tabulations/calculations the 9-filled values will be treated as any other numeric data. Prior to calculating univariates or any other tabulations for variables defined as numeric, remove all 9-filled values from the universe to prevent calculation skewing which would produce inaccurate results.

Data users must define the appropriate universe definition for each variable.

#### *2. Demographic Variables - Technical Notes*

The specific notes below address standards for editing relationship to reference person and cases that switch values in the unedited data for sex, vet, and race status. Note that in 1997, using the CPS instrument, Field Representatives could change demographic data that was input since the data provided was from SIPP paper control cards. If the information was incorrect during SIPP, they got a chance to change it in 1997, yet because demographic variables are edited forward from 1994, the edits force the responses to be consistent with the original SIPP data. That is why there is a higher number of imputations for 1997 than for other years.

##### *a. Relationship to Reference Person (RRP)*

An adult must be at least 12 years older than a child who points to them as their designated biological parent.

Same sex partners who report they are married are made unmarried partners if they report being spouses.

People less than 15 years old must be 'never married.'

A person to whom a child points as their designated parent must be at least 15 years old.

A person under 18 years of age cannot be the reference person if their parent is present in the household.

In 1999 through 2002, we edited and output the designated parent pointer for people of all ages, so checks for the age gap between parents and children were done for people of all ages. For 2000, we also used the TYPMOM and TYPDAD variables which indicate whether a parent is a biological, step or adoptive parent. If the parent was identified as a step or adoptive parent, we allowed an otherwise illogical age gap between parent and child since remarriage sometimes creates situations where step parents are not old enough to be the biological parent of their spouse's biological children. Users should note that the edit for RRP, MARITL, SPOUSE, and DESPAR is not truly longitudinal, but checks for consistency in each year cross sectionally. So the RRP values of two people shift over time, even though they are the same people. This happens at times when a later interview gives us more information and we are able to specify the relationship in greater detail—for example, making someone a grandchild rather than other relative.

#### ***b. Age***

We have not filled in DOBYRT for all cases, although users could easily create it using the year and the age. At this point it is filled in for interviews in 1998 and 1999 and 2000.

#### ***c. Sex***

Demographic variables are edited forward in the SPD. So the earliest value for SEX is carried through the entire time the person is interviewed. The imputation variable SEXI (if it =4) will let you know which cases contained contradictory reports over time in the unedited data. Cases which have a record in the 2<sup>nd</sup> longitudinal file were kept the same sex as they were in that file if they reported being a different sex in the 3<sup>rd</sup> longitudinal file unedited data.

#### ***d. Race***

Each person receives one value for RACE, since this is not allowed to change over time. We take the earliest valid value and hold it constant. So for people who changed their reported RACE, we keep the earliest reported race regardless. Note that since we do not allow a value of "other" in the output file, new people who reported "other" are imputed into one of the listed race groups. For people who were already on the file previously, and so have a RACE value, but who report a value different from the one which they have on the SPD 1<sup>st</sup> longitudinal file, we note this by making the value of RACEI=4 for these cases.

#### ***e. NATVTY***

We have added a value of 4 for the imputation flag so that we can indicate cases whose value on the first longitudinal file differs from the value they reported in an earlier year.

### *f. Family*

The family is defined by the Census Bureau as two or more people living together who are related by blood, marriage, or adoption. On this file, it is possible to distinguish primary families (the family containing the household reference person) from related subfamilies and unrelated subfamilies. A related subfamily is either a parent-child(ren) group or married couple who are related to the household reference person. An unrelated subfamily is a parent-child(ren) group or married couple that does not contain the household reference person or anyone related to the household reference person.

### *g. Origin*

Although ORIGIN has been weighted, these data will provide very different estimates of some specific Hispanic subgroups than CPS 1999 data. So these data should probably not be used for analysis of particular Hispanic groups.

## **3. Income Variables**

### *a. Creating spells from the minimally edited data*

Data users must create spells in order to define universes for several income source variables. A spell is the length of time a respondent received an income source. Since some income sources can stop and start throughout the year, the survey allows for up to six individual spells over the course of a year. In reviewing the unedited data at most three spells were consistently reported for all income sources and, hence, only data for up to three spells are provided.

There are four questions for each spell:

- Part A:** What set of circumstances led you to apply?;
- Part B:** Why did you stop receiving the income source?;
- Part C:** What reasons were given for your benefits being cut off?;
- Part D:** What did you do to get by when your family lost benefits?

These questions, Parts A through D, are in reference to when there are distinct starting and stopping points for a time period of interest during a calendar year, rather than to the length or duration of a spell of participation. The following examples provide insight into how spells can be created using the minimally edited 2000 SPD cross-sectional data, as well as when the above questions, Parts A through D, would have been asked in the survey.

It is the convention in longitudinal analysis to construct a spell when there are distinct stopping and starting points, that is you have both right and left censoring. In using the minimally edited 2000 SPD cross-sectional data, to define the first spell, you will need to review the monthly receipt of each income source. For example, if someone said they received Supplemental Security Income (SSI) from February to March (SSIM02, SSIM03 = X) and from October to November (SSIM10, SSIM11=X) and all other months are equal to a blank, then spell one is February to March and spell two is October to November. In this example there is no third spell. For each of these time periods there are distinct starting and stopping points and therefore questions from Parts A through D would have been asked for each spell. However, if someone reported receiving SSI for all months, SSIM01-SSIM12 = X then Parts A through D are not asked, since there is no spell, just a continuing program participation.

Next, consider the situation where there is not a distinct starting point, but a distinct stopping point. Suppose the respondent said they received a source of income, such as SSI in January, February and March (SSIM01, SSIM02, SSIM03 = X) of the year. Similarly, suppose they reported receiving SSI from October to November (SSIM10, SSIM11 = X). In this situation, the user will need to make further clarification for the universe of respondents to Parts A through D. The first time period, January through March, does not have a distinct starting point and therefore can not be considered a spell. It is unclear when they started to receive this source of income, but it is clear when they stopped receiving it. They could have started receiving it some time in the prior year or years. We don't know when they started to receive, only that they were receiving this income in January. On the other hand, for the second time period, October to November, there is a distinct starting and stopping time, so we have a defined spell. For the first time period, January through March, Part A is not asked because there is not a distinct starting point. However, questions from Parts B, C and D are asked, since there is a distinct stopping time. For the time period October to November, there is a distinct starting and stopping point, which would count as a spell for the calendar year, therefore questions from Parts A through D would be asked for this spell.

Finally, consider the situation when the respondent said they received the source of income in January and December of the same year. If someone received SSI from January to March (SSIM01, SSIM02, SSIM03 = X) and from November to December (SSIM11, SSIM12= X), Part A questions would apply only to the time period covering November to December. Why? Because there is a distinct starting time, November, but not a distinct ending time. Since there was no distinct stopping time for this time period, Part B through D would not apply. On the other hand, for the January to March time period, since there is no distinct starting time, but a distinct stopping time, Part A questions would not apply, but Parts B through D questions would be applicable.

***b. Unemployment Compensation (V201) and Social Security payment on behalf of a child (V209) variables***

UCYNU0 - Did you receive unemployment compensation payments at any time during 1999?, Yes/No (V201)

SSCHYNU0 -Did you receive any separate Social Security payment on behalf of a child? Yes/No (V209)

In the process of preparing the 2000 minimally edited cross-sectional data, an error in the data collection instrument was uncovered. In the process of collecting the data, if a respondent said "No" to question V209, 'Did they receive Social Security payments on behalf of a minor or dependent child in the household', the value for V201 was set to a "2" rather than the value of V209 being set to a "2". Hence, the minimally edited cross-sectional data for 2000 will be more negatively biased in regard to responses for V201 and be more positively biased in the responses reported for V209. The reported values for UCYNU0 and SSCHYNU0 should be used with caution and other minimally edited variables should be used, whenever possible, to investigate issues related to receipt of unemployment compensation and receipt of social security payments on behalf of a minor or dependent child.

***c. Spells of Program Participation - Variables with "No Values" reported.***

While it is possible to determine what type of cash assistance and individual received from their responses to Z221A\_ADD1 - Z221A\_ADD4, the data instrument only collected data on the total amount of cash assistance received for the calendar year. This also holds true for questions dealing with the duration of receipt of the cash assistance, as well as other questions dealing with

applying for assistance, the reasons they are currently not receiving cash assistance, loss of cash assistance and how they got by after losing cash assistance. Because of this error the following variables in the minimally edited 2000 SPD cross-sectional file have “no values” reported.

OA1A01U0, OA1A03U0-OA1A10U0 -- why applied for other assistance, spell 1  
OA2A01U0, OA2A03U0-OA2A09U0 -- why applied for other assistance, spell 2  
OA3A01U0-OA3A10U0 -- why applied for other assistance, spell 3  
OA2STPU0 -- Reason stopped other assistance, spell 2  
OA3STPU0 -- Reason stopped other assistance, spell 3  
OA1C01U0-OA1C13U0 -- Reason cut-off other assistance, spell 1  
OA2C01U0-OA2C13U0 -- Reason cut-off other assistance, spell 2  
OA3C01U0-OA3C13U0 -- Reason cut-off other assistance, spell 3  
OA1D01U0-OA1D13U0 -- How got by when cut-off other assistance, spell 1  
OA2D01U0-OA2D13U0 -- How got by when cut-off other assistance, spell 2  
OA3D01U0-OA3D13U0-- How got by when cut-off other assistance, spell 3

The following variables in the minimally edited 2000 cross-sectional file have “no values” reported. This is due to the nature of the question, but more importantly, the nature of the program, WIC, and any spells of program participation in calendar year 1999. No one reported three spells of WIC participation during the year.

WC3D01U0-WC3D13U0 -- How got by when cut-off from WIC, spell 3

#### ***4. General note - Child and Family Well-Being Variables***

The SPD 2000 Cross-sectional File is minimally edited. The child and family well-being sections are not edited. Users should take caution to properly define the universe for each item when using the data.