

APPENDIX F

User Notes for the SPD 2001 Minimally Edited Cross-sectional File

This section contains information relevant to the Survey of Program Dynamics 2001 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 "Publications and Analyses". In addition, some user notes are found under "Data Access". The Internet site will be updated as additional user notes become available.

User Notes for the SPD 2001 Minimally Edited Cross-sectional File

1. *General Notes*

The data on the SPD 2001 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.

Blanks appear on the file and are defined as one of the following:

- Not in Universe
- No response reported
- No (for mark all that apply responses)

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

2. *Demographic Variables*

a. Educational Attainment Variables

Individuals age 18 and over who indicated in response to EDUCA that they were NOT high school graduates (EDUCA = 31-38) were asked whether they had completed high school by means of a GED. The 191 individuals who responded "Yes" had their values of EDUCAU1 set to 39, high school graduate-high school diploma or equivalent (e.g., GED).

b. 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.

c. *Weighted estimates by age*

Although AGE has been edited, the weighted estimates yield lower numbers of 0 to 4 year olds, and higher numbers of 10 to 14 year olds than CPS 2001 data, so any analysis of these age groups is discouraged. Caution should be used regarding analysis of the population 0 to 17.

d. *Parent variables*

Parent Pointers:

DESPARE1 is edited for all people under 18 more thoroughly with respect to LNMOM and LNDAD than for people 18 and over. That is, if there was no designated parent reported, for people under 18, we look to see if they point to a mother or father who is in the household, and make that person their designated parent. For people 18 and over, if they report a designated parent, we edit that relationship for consistency, but do not look further to try to identify their parents.

Type of Parent:

Thousands of people have valid values for TYPMOM (4,668) or TYPDAD (3,026) although LNMOM or LNDAD does not point to their parent. Some of these may be out of universe—that is, their parent may not be in the household. However, for most, it appears that LNMOM or LNDAD is not filled in and simply doesn't point to their parent, although their parent is present in the household. Users may find some of these parents based on the edited RRP values. Eighty-six percent of those who have a valid TYPMOM, but do not point to a mom in the household, based on LNMOM, are the children of the reference person. Eighty-nine percent of those who have a valid TYPDAD, but do not point to a dad in the household, based on LNDAD, are the children of the reference person. Users are cautioned to use this information in order for analysis of the type of relationship between parents and children, since most of these cases are people under 18.

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 2001 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 2001 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

UCYNU1 - Did you receive unemployment compensation payments at any time during 2000?, Yes/No (V201)

SSCHYNU1 -Did you receive any separate Social Security payment on behalf of a child? Yes/No (V209)

In the process of preparing the 2001 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 2001 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 UCYNU1 and SSCHYNU1 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. Variables with “No Values” reported.

Because of an error on the data collection instrument, the following variables in the minimally edited 2001 SPD Cross-sectional file have “no values” reported. While it is possible to determine what type of cash assistance an individual received from their responses to the variables CAWELFUL1, CAGENU1, CAEMU1 AND CAOTHRU1 the 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. The variable affected by this problem are:

GAM01U1 - GAM12U1 -- Month received general assistance in 2000
GAMTU1 - Total number of months received general assistance in 2000
OWEM01U1 - OWEM12U1 -- Month received other assistance in 2000
OWEMSU1 - Total number of months received other assistance in 2000
OA1A01U1-OA1A10U1 -- why applied for other assistance, spell 1
OA2A01U1-OA2A10U1 -- why applied for other assistance, spell 2
OA3A01U1-OA3A10U1 -- why applied for other assistance, spell 3
OA1STPU1 -- Reason stopped other assistance, spell 1
OA2STPU1 -- Reason stopped other assistance, spell 2
OA3STPU1 -- Reason stopped other assistance, spell 3
OA1C01U1-OA1C13U1 -- Reason cut-off other assistance, spell 1
OA2C01U1-OA2C13U1 -- Reason cut-off other assistance, spell 2
OA3C01U1-OA3C13U1 -- Reason cut-off other assistance, spell 3
OA1D01U1-OA1D13U1 -- How got by when cut-off other assistance, spell 1
OA2D01U1-OA2D13U1 -- How got by when cut-off other assistance, spell 2
OA3D01U1-OA3D13U1 -- How got by when cut-off other assistance, spell 3
OAVALT1 - Amount of other welfare cash assistance received
GASVALT1 - Amount of general cash assistance received
EMGVALT1 - Amount of emergency assistance received

The following variables in the minimally edited 2001 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 2000. No one reported three spells of WIC participation during the year.

WC3D01U1-WC3D13U1 -- How got by when cut-off from WIC, spell 3

4. Household Food Security Variables

Due to a problem with the instrument 548 individuals, most from households containing only 1 child under 18 years of age, were mistakenly skipped over FOOD04U1 and FOOD05U1 and 522 individuals were mistakenly skipped over FOOD06U1.

Due to a problem with the instrument 1,495 individuals living in households without children were mistakenly skipped over FDBANKU1, FDBFRQU1, and FDSOUPU1.

5. Child School Enrollment Variables

a. Enrolled in preschool

There are 100 individuals who are out of universe for the preschool variables (PRESCHU1-PREHRU1) but have valid values for these variables.

There are 333 individuals who are in-universe for PRESCHU1 but do not have a valid value for this variable. In 38 of these cases the individuals DO have valid values for one or more of the monthly enrollment variables, PREM06U1-PREM17U1.

There are 194 individuals who indicated they were not enrolled in preschool (PRESCHU1=2) but have valid values for one or more of the monthly enrollment variables, PREM06U1-PREM17U1. There are 38 individuals with missing values for PRESCHU1 and valid values for one or more of the monthly enrollment variables, PREM06U1-PREM17U1. There are two individuals with values of 'D' or 'R' for PRESCHU1 and valid values for one or more of the monthly enrollment variables, PREM06U1-PREM17U1.

b. Enrolled in regular school

There are 44 individuals who are out of universe for the regular school enrollment variables (SCHOOLU1-SCWHYU1) but have at least some valid values for these variables.

There are 1,369 individuals who are in universe for SCHOOLU1 but do not have a valid value for this variable. In 1,000 of these cases the individuals have no school enrollment data. In 366 of these cases the individuals DO have valid values for one or more of the monthly enrollment variables, SCHM06U1-SCHM17U1. In 3 of these cases the individuals have missing data for SCHOOLU1 and SCHM06U1-SCHM17U1 but valid data for other regular school enrollment variables.

There are 405 individuals who have valid values for one or more of the monthly enrollment variables, SCHM06U1-SCHM17U1, even though they were not enrolled in regular school (SCHOOLU1 ne 1). Of these, 366 have missing data for SCHOOLU1, 33 have SCHOOLU1=2, and 6 are not-in-universe for the regular school variables.

6. Child Care Variable

R01WREU1 - Where child cared for most of the time.

Sixteen unweighted cases are coded "4." The value "4" is invalid, and the cases should be recoded to missing by the analyst.