

PANEL STUDY OF INCOME DYNAMICS
1996 EARLY-RELEASE FILES
DOCUMENTATION

TABLE OF CONTENTS

I. INTRODUCTION

A. REASONS FOR AND LIMITATIONS OF THE EARLY-RELEASE FILES

B. WHAT'S NEW FOR 1996

1. HOME FINANCING QUESTIONS
2. FINANCIAL DISTRESS QUESTIONS
3. RISK AVERSION QUESTIONS

II. CHARACTERISTICS

A. FILES AND FORMAT

1. .DAT FILES
2. .SAS AND .SPS FILES
3. .TXT FILES

B. VARIABLE NUMBERS, POSITIONS AND GENERATED VARIABLES

1. FAMILY FILE VARIABLES
2. CROSS-YEAR INDIVIDUAL FILE VARIABLES

C. DOCUMENTATION AND CODES

1. 1996 QUESTIONNAIRE
2. CODE CATEGORIES

D. BRINGING FORWARD BACKGROUND INFORMATION FOR HEAD AND WIFE/"WIFE"

1. ADD 1996 HEAD'S 1992-1995 INTERVIEW NUMBERS TO INDIVIDUAL FILE
2. OBTAIN INFORMATION FROM 1996 FAMILY FILE
3. OBTAIN INFORMATION FROM 1995 FAMILY FILE
4. OBTAIN INFORMATION FROM 1994 FAMILY FILE
5. OBTAIN INFORMATION FROM 1993 FAMILY FILE
6. OBTAIN INFORMATION FROM 1992 FAMILY FILE

E. PROBLEM VARIABLES, MISSING VARIABLES

1. FAMILY FILE PROBLEM VARIABLES
2. CROSS-YEAR INDIVIDUAL FILE PROBLEM VARIABLES
3. CHANGE IN FAMILY-LEVEL VARIABLE LABELS

F. COMPARABILITY WITH 1992 FINAL-RELEASE FILE AND EARLIER FINAL-RELEASE FILES

G. ADDITIONAL NOTES: SAMPLE SUPPLEMENTS IN 1993, 1994, 1995 AND 1996

I. INTRODUCTION

A. Reasons for and Limitations of the Early-Release Files

The more than two-year interval between the completion of interviewing on a given PSID wave and the public release of a fully cleaned and documented data file has prompted demand for speedier release of an "early-release" version of PSID data files. In response to this demand, the PSID staff produced an early-release version of the 1990, 1991 and 1992 family files and the 1968-1992 cross-year individual file; all these files are now available in final-release form.

These were followed by early-release versions of the 1993, 1994, 1995 and 1996 waves of data. The latest in this series is the early-release version of the 1996 wave of data -- an early-release version of the 1996 family file and of the 1968-1996 cross-year individual file. The early-release files are available both at the PSID's Internet site and from the ICPSR (Inter-university Consortium for Political and Social Research).

The advantage of the early-release files is that of quicker access to recent waves of data; disadvantages include:

- * more limited documentation -- for the 1996 early-release files, no documentation other than: this readme file, univariate frequencies of categorical variables, univariate statistics for continuous variables, a cross-year index for the 1993, 1994, 1995 and 1996 early-release files, and the 1996 questionnaire;
- * missing variables -- a set of variables that differs substantially from that provided for 1992 and earlier years, e.g., you will find two variables, dollar amount and periodicity, in place of one variable, dollars per hour and an incomplete set of family and individual variables, with the most prominent missing variables in the family file being the annualized work and income components and totals including total family income, generated variables such as the income to needs ratio and prorated poverty thresholds, open-ended variables such as occupation and industry, and sampling weights and in the individual file the "summary variables" including variables about marital and fertility histories;
- * minor problems with data values -- no imputations for missing data; data containing some wild codes; scrambled data for some variables in a few cases where a different person was determined to be the correct Head or Wife/"Wife" during family composition editing; zero values for most or all cases for a handful of variables (see detailed description of these limitations in Section II below); and

B. What's New for 1996

1. Home Financing Questions

We added questions about attempted home financing. Homeowners with a mortgage or other loan on their main residence were asked the type of loan, whether it was insured by a government agency, whether it was the original loan or refinanced, the type of interest rate and the actual rate. If the loan was obtained in 1991 or later, more details of the financing experience were asked, such as the source for a down payment; what kind of relationship, if any, the homeowner had had with the lender; how the homeowner had come to know the lender, and whether any other loan options had been considered. If they had, reasons for failure to use these sources were explored.

Homeowners with mortgages who had last obtained financing on the current dwelling before 1991 were asked whether they had been turned down for financing from 1991 onward, with reasons for financing refusal.

All non-homeowning respondents were asked whether they had attempted to purchase a home since 1991. If they had, questions detailing the reasons for the non-purchase were asked. See Section A of the 1996 questionnaire for actual question wordings.

2. Financial Distress Questions

For the 1996 wave only, we added a lengthy series about financial distress, located at the end of Section G, Income. The distress questions are numbered from G115 on and were asked after all of the usual annual income questions had been asked. This supplement includes information about consolidation loans, creditors demanding payment, wage attachment, and liens against and repossession of property. For each such instance, questions about up to two of the most recent instances of each type of financial problem were asked. Specific questions include the kinds of debts held; the length and amount of garnishments; and the kinds of property repossessed or against which a lien was taken.

As part of this series, data were gathered about as many as two bankruptcy filings, with details about the reasons for filing and court-ordered repayment plans. Post-bankruptcy credit implications were also queried. See questions G115-G146 for

exact question wordings and specific details.

3. Risk Aversion Questions

Five questions about risk aversion were asked at the end of the interview. See questionnaire Section M for the text of these questions.

II. CHARACTERISTICS

A. Files and Format

The early-release package for 1996 wave consists of two data files: the 1996 early-release family data file, and the 1968-1996 early-release individual data file.

File name	Number of variables	Number of Records	LRECL	Bytes	Contents
ER96F.DAT	2,107	8,511	3,585	30,528,957	1996 raw family data
ER68-96I.DAT	935	57,409	2,005	115,219,863	1968-1996 raw individual data

In addition to these two data files and the file you are reading, ER96READ.ME, the 1996 early-release files include a number of other files listed in the tables below. The contents of these files are described in more detail paragraphs below.

Other 1996 Early-Release Family Files

File name	Bytes	Contents
ER96F.SAS	106,428	SAS statements for family variables
ER96F.SPS	106,319	SPSS statements for family variables
ER96FMEA.TXT	140,685	means for family variables
ER96FTAB.TXT	525,972	frequencies for family variables
ER96FMD.TXT	102,539	SAS missing data statements

Other 1968-1996 Early-Release Individual Files

File name	Bytes	Contents
ER68-96I.SAS	62,331	SAS statements for individual variables
ER68-96I.SPS	60,159	SPSS statements for individual variables
ER96ITAB.TXT	182,462	frequencies and means for individual variables

1. .DAT files

The data are in raw ASCII form. Refer to one of the corresponding files -- .SAS or .SPS -- for record format layout information, variable names and variable labels.

The 1996 early-release family data file contains one record for each family interviewed in 1996. The file includes all family-level variables collected in 1996.

The 1968-1996 early-release individual data file is a merged cross-year file that contains a record for both 1996 response and 1996 nonresponse individuals; it includes a record for each individual who was in an interviewed family for any wave of the study. The file includes individual-level variables collected from 1968 through 1996.

2. .SAS and .SPS files

These files, respectively, contain SAS and SPSS data definition statements which provide variable names, locations, and variable labels. The SPSS and SAS statements are NOT intended to represent completed and full programs for the respective statistical program packages to run extracts, analysis, etc. You

must provide all other SPSS or SAS statements needed to complete a program.

Missing data statements have not been provided. You should check the questionnaire and frequencies or means for each variable you intend to include in your analysis to determine which code values should be defined as missing. We do plan to include traditional missing data information with the final-release versions.

3. .TXT files

These ASCII text files provide additional information about the early-release family file.

The file ER96FTAB.TXT contains univariate frequencies for the categorical family variables. The file ER96ITAB.TXT contains univariate frequencies for the early-release categorical individual variables and univariate statistics for the few early-release continuous individual variables. These files may be used to check for wild codes.

The file ER96FMEA.TXT contains univariate statistics for the continuous family variables. For your information, the ad-hoc missing data statements used in producing these means are included at the end of the file. The missing data statements for the final-release file will differ from these ad-hoc missing data statements because these statements were modified to include as missing some inappropriately entered values.

1. Family File Variables

The 1996 early-release family variables are in the range ER7001-ER9107. Most of these variables will eventually be incorporated into the final version of the 1996 data, but their variable numbers will change and the data will be cleaner.

Variable numbers and locations for the 1996 early-release family file are not the same as those we intend for the final version. The 1996 early-release family file includes neither variable numbers nor positions for so-called "edited" and "generated" family-level variables.

* By "edited" variables we mean the first 300 or so variables usually present in each wave's family-level data, beginning with the state of residence and ending with income detail for other family unit members.

* By "generated" variables we mean those variables traditionally located at the end of the raw data after the Head's background information.

In short, all variables equivalent to the 1992 variable ranges V20303-V20620 and V21481-V21549 are absent.

Variables not included in the early-release file for which component items are available include:

- * annual mortgage and rent payments,
- * annual food costs,
- * poverty thresholds,
- * annual work hours,
- * annual unemployment, etc., hours,
- * annual income of any sort for Head and Wife/"Wife",
- * Head's total labor income,
- * numbers of children in various age and sex categories,
- * education of Head and Wife/"Wife", and
- * average hourly earnings of Head and Wife/"Wife".

Since component items exist on the early-release file, you may generate these items. Imputations have NOT been done for missing data. To create variables from the 1996 early-release data that resemble those on final files from 1992 and earlier waves, we suggest you consult the 1992 codebooks where you will find sufficient information about how the variables were created for 1992 to create them for 1996.

Background information has not been asked about all Heads and Wife/"Wives" each and every year. We ask the questions for new Heads and new Wives/"Wives" only. During processing, we have traditionally "brought forward" the background information from previous waves for Heads or Wives/"Wives" who are the same persons as in the prior year. In every wave, each set of background variables is preceded by a variable indicating whether data needed to be brought forward. The 1996 early-release file, in keeping with our practice for other early-release files, has not undergone this "bringing forward". See Section D below for a detailed description of how you can do this yourself.

Other variables are not generatable because income components of individuals other than Head and Wife/"Wife" are not included in the 1996 early-release data. Variables not included in the early-release which cannot be generated from available information include:

- * annual income of any sort for other family members,
- * total family money income,
- * poverty thresholds (because of missing income components),
- * family income deciles,
- * sampling weights,
- * state and region of residence,
- * urbanicity,
- * Head's geographic mobility,
- * county unemployment rate, and
- * variables linking related families.

2. Cross-Year Individual File Variables

Recent cross-year individual PSID files have consisted of annual measures and a set of "summary variables" that have appeared at the end of the individual data record. In the 1968-1996 early-release individual file, most of the annual measures (e.g., Sequence Number, Relationship to Head, Family Identification Numbers) are available. However, virtually NONE of the "summary variables" (i.e., V31996-V32049) are included; the single exception is V32000, Sex of Individual, which was too important to omit; it appears in the 1968-1996 early-release individual file as ER32000.

Variables ER30001 through ER30794 will remain the same for the final-release version (with the prefix change from "ER" to "V"). A few more variables will be added to the 1992-1996 individual data, most notably the sampling weights. The order of variables in the 1968-1996 early-release individual file is as follows:

- * RELEASE NUMBER, ER30000,
- * 1968 through 1992 individual data arranged, as usual, by wave, ER30001-ER30794
- * the lone summary variable, SEX OF INDIVIDUAL, ER32000,
- * the 1993 early-release individual data, ER33001-ER33018,
- * the 1994 early-release individual data, ER33101-ER33118,
- * the 1995 early-release individual data in ER33201-ER33274,

and

- * the 1996 early-release individual data in ER33301-ER33018.

For the final-release version, the 1993, 1994, 1995 and 1996 variables will be moved to follow the completed 1992 individual data and ER32000 will appear in its usual place among the summary variables as V32000.

Some 1993, 1994, 1995 and 1996 equivalents of traditional annual individual variables are not included in the early-release file:

- * individual income components and totals,
- * linking measures for splitoffs,
- * reason for nonresponse, and
- * sampling weights.

In the final-release individual files, these variables will

be located near the end of the yearly data, just as in 1992 and earlier waves.

C. Documentation and Codes

1. 1996 Questionnaire

We have not produced the traditional codebooks for the early-release files. A 1996 questionnaire is available at our site on the Internet in a PDF format suitable for perusal with an Adobe Acrobat viewer (the Acrobat viewer is available free of charge - see our home page for further information). Use the SAS and SPSS data definition statements to match variables with questions in the questionnaire. The questionnaire contains codes for most data items.

2. Code Categories

In addition to the 1996 questionnaire, for family data, the codebook from Section II, Part 1 of the 1992 documentation can also be helpful in deciphering the early-release data. For individual data, use the codebook in our 1992 documentation, Section II, Part 2; similar variables for 1993, 1994, 1995 and 1996 are coded identically to those from earlier waves.

In general, codes follow our traditional structure, although "don't know" responses are now largely distinguished from other missing data responses. If the questionnaire does not indicate otherwise, code 8 (or 98 or 998, etc.) represents "don't know" and code 9 (or 99 or 999, etc.) represents other missing data or a refusal. Inappropriate questions are padded with zeroes. A few fields contained non-numeric characters, and these have also been converted to zeros for the early-release file. If a variable contains a code value that is neither included in the questionnaire nor one of the zero, eight or nine codes just mentioned, you should assume missing data for that value. We will clean such cases for final-release, but time constraints do not permit this sort of cleaning for early-release.

The inevitable exception: codes 21 through 24 for month variables in event dating questions were not printed in the questionnaire but were used throughout the CATI application to indicate mentions of season only. These codes follow:

- 21. DK month, but season was winter
- 22. DK month, but season was spring
- 23. DK month, but season was summer
- 24. DK month, but season was autumn

D. Bringing forward Background Information for Head and Wife/"Wife"

As noted above, the background information for Head and Wife/"Wife" has not been "brought forward" for the 1996 early-release family file. Background information is complete for 1992 on the 1992 final-release family file, but as of this writing, the 1993, 1994 and 1995 family data are available only in early-release form and have not yet undergone the bringing-forward process. Only families with Heads and Wives/"Wives" who were new in 1996 have background data in the 1996 early-release family file. You must search, respectively, the 1995, 1994, 1993 and 1992 family data to complete 1996 background variables.

Carefully compare the background variables item for item and code for code in 1992 final-release family file and the 1993, 1994, 1995 and 1996 early-release family files before you attempt to bring forward prior-wave background information. You should be aware that the 1993, 1994, 1995 and 1996 background variables are not necessarily completely identical to each other!

In addition, some 1992 background questions are not included at all in the 1993, 1994, 1995 and 1996 early-release family files' background data because they have NOT YET been created; among these are questions about:

- * Head's father's occupation,
- * state and county variables for the locations where Head and

his/her parents grew up, and

* number of states and regions in which Head has lived.

One more factor complicates bringing forward background data: the absence of the 1992, 1993, 1994 and 1995 family ID numbers on the 1996 family file. You must obtain these variables from the Head's record in the 1968-1996 early-release individual file in order to match with 1992, 1993, 1994 and 1995 family files to bring forward the background information.

Below you will find detailed a suggested procedure for bringing forward the Head's and Wife's/"Wife's" background information.

1. Add 1996 Head's 1992-1995 Interview Numbers to Individual File

First, add the 1996 Head's 1992, 1993, 1994 and 1995 interview numbers from the 1968-1996 early-release individual file to the 1996 early-release family file.

```
sort er96fam by 96 i'w (ER7002 "1996 INTERVIEW #")
```

```
sort er68-96ind by 96 i'w (ER33301 "1996 INTERVIEW NUMBER")  
for 1996 Heads (ER33302 "INDIVIDUAL SEQUENCE NUMBER 96" =  
01)
```

```
merge er96fam & er68-96ind by 96 i'w; add Head's 1995, 1994,  
1993 and 1992 i'w # (ER33201 "1995 INTERVIEW NUMBER",  
ER33101 "1994 INTERVIEW NUMBER 94", ER33001 "1993 INTERVIEW  
NUMBER 93", ER30733 "1992 INTERVIEW NUMBER") to er96fam
```

2. Obtain Information from 1996 Family File

Check to determine whether the 1996 family includes a Wife/"Wife" and whether new Head and new Wife/"Wife" information is present in the 1996 early-release family file. If it is, then the appropriate background information is already part of the 1996 early-release family file, and this case needs no further processing.

```
if no Wife/"Wife" (ER7008 "AGE OF WIFE" = 0)then statwife=1  
else statwife=0
```

```
if new Wife/"Wife" 96 (ER8979 "K1 CKPT: WTR WIFE" = 1)then  
statwife=1
```

```
if new Head 95 (ER9033 "L1 CKPT: WTR NEW HEAD" = 1)then  
stathead=1 else stathead=0
```

3. Obtain Information from 1995 Family File

If new Head or new Wife/"Wife" information was not present in the 1996 early-release family file, check to determine whether it is present in the 1995 early-release family file. If it is, then replace the values of the variables in the 1996 early-release family file with values of the corresponding variables from the 1995 early-release family file. Remember that these variables differ slightly from year to year.

```
sort er96fam by 95 i'w (ER33201 "1995 INTERVIEW NUMBER 95")
```

```
sort er95fam by 95 i'w (ER5002 "1995 INTERVIEW #") for new  
95 Wife/"Wife"s (ER6733 "K1 CKPT: WTR WIFE" = 1) or for new  
95 Heads (ER6787 "L1 CKPT: WTR NEW HEAD" = 1)
```

```
merge er96fam & er95fam by 95 i'w
```

```
if statwife=0 and ER6733=1, bring forward 95 new  
Wife/"Wife" info and set statwife=1
```

```
if stathead=0 and ER6787=1, bring forward 95 new Head  
info and set stathead=1
```

4. Obtain Information from 1994 Family File

If new Head or new Wife/"Wife" information was not present in either the 1996 nor in the 1995 early-release family files, check to determine whether it is present in the 1994 early-release family file. If it is, then replace the values of the variables in the 1996 early-release family file with values of the corresponding variables from the 1994 early-release family file. Again, recall that differ slightly from year to year.

```
sort er96fam by 94 i'w (ER33101 "1994 INTERVIEW NUMBER 94")
```

```
sort er94fam by 94 i'w (ER2002 "1994 INTERVIEW # 94") for  
new 94 Wife/"Wife"s (ER3863 "K1 CKPT: WTR WIFE" = 1) or for  
new 94 Heads (ER3917 "L1 CKPT: WTR NEW HEAD" = 1)
```

```
merge er96fam & er94fam by 94 i'w
```

```
if statwife=0 and ER3863=1, bring forward 94 new  
Wife/"Wife" info and set statwife=1
```

```
if stathead=0 and ER3917=1, bring forward 94 new Head  
info and set stathead=1
```

5. Obtain Information from 1993 Family File

If new Head or new Wife/"Wife" information was not present in either the 1996 nor in the 1995 or 1994 early-release family files, check to determine whether it is present in the 1993 early-release family file. If it is, then replace the values of the variables in the 1996 early-release family file with values of the corresponding variables from the 1993 early-release family file. Again, recall that differ slightly from year to year.

```
sort er96fam by 93 i'w (ER33001 "1993 INTERVIEW NUMBER 93")
```

```
sort er93fam by 93 i'w (ER2 "1993 INTERVIEW # 93") for new  
93 Wife/"Wife"s (ER1777 "K1 CKPT: WTR WIFE" = 1) or for new  
93 Heads (ER1850 "L1 CKPT: WTR NEW HEAD" = 1)
```

```
merge er96fam & er93fam by 93 i'w
```

```
if statwife=0 and ER1777=1, bring forward 93 new  
Wife/"Wife" info and set statwife=1
```

```
if stathead=0 and ER1850=1, bring forward 93 new Head  
info and set stathead=1
```

6. Obtain Information from 1992 Family File

If new Head or new Wife/"Wife" information was not present in the 1996, 1995, 1994 or 1993 the early-release family files, obtain the information from the 1992 final-release family file. There is no need to check the value for the 1992 indicator, as all 1992 cases contain background information. Replace the values of the variables in the 1996 early-release family file with values of the corresponding variables from the 1992 final-release family file. Again, recall that these variables do not match perfectly. And you are done. Congratulations.

```
sort er96fam by 92 i'w (ER30733 "1992 INTERVIEW NUMBER")
```

```
sort 92fam by 92 i'w (V20302 "1992 INTERVIEW NUMBER")
```

```
merge er96fam & 92fam by 92 i'w
```

```
if statwife=0, bring forward 92 Wife/"Wife" info and  
set statwife=1
```

```
if stathead=0, bring forward 92 Head info and set  
stathead=1
```

E. Problem Variables, Missing Variables

Some variables included on the 1996 early-release files are known to include bad or completely missing data. These will be corrected for the final version of the file, but in the meantime we want you to be informed of the following known problems with

the early-release data.

1. Family File Problem Variables

Some family information included in prior early release files is missing in 1996. For Heads, the three mentions (questions B21 and C2, ER7190-ER7192 and ER7413-ER7415) of the types of actions taken to find new employment are not available for the 1996 early release file. The equivalents for Wives/"Wives", D21 and E2 (ER7684-ER7686 and ER7907-ER7909), are also missing. The questions were, however, asked in 1996, and the data will be included in the 1996 final release dataset. Values for the following variables are also all blank in the early-release version of the dataset.

ER7002	1996 INTERVIEW #
ER7253	B52 BEG HR/WK OTR EMP HD
ER7498	C44 BEG HR/WK OTR EMP HD
ER7747	D52 BEG HR/WK OTR EMP WF
ER7992	E44 BEG HR/WK OTR EMP WF

The year in which Head or Wife/"Wife" began an extra job is also unfortunately blank. This affects as many as four data items for each Head or each Wife/"Wife", as we include information about up to four extra jobs held in the prior calendar year. For employed Heads, the variables are:

ER7330	B90 YR BEG XTRA JOB1 H-E
ER7352	B102 YR BEG XJOB2 (H-E)
ER7372	B113 HR/WK XTRA JOB3
ER7396	B126 YR BEG XTRA JOB4

If Head is currently not working, the potentially blank variables are:

ER7575	C82 YR BEG XTRA JOB1 H-U
ER7597	C94 YR BEG XJOB2 (H-U)
ER7619	C106 YR BEG XTRA JOB3
ER7641	C118 YR BEG XTRA JOB4

The equivalents for Wives/"Wives" are:

ER7824	D90 YR BEG XTRA JOB1 W-E
ER7846	D102 YR BEG XJOB2 (W-E)
ER7868	D114 YR BEG XTRA JOB3
ER7890	D126 YR BEG XTRA JOB4
ER8069	E82 YR BEG XTRA JOB1 W-U
ER8091	E94 YR BEG XJOB2 (W-U)
ER8113	E106 YR BEG XTRA JOB3
ER8135	E118 YR BEG XTRA JOB4

ER7317 "B78 # WKS WORKED (HD-E)" and ER7319 "B COMPUTED WKS WORKED" should, conceptually, have identical values but, because of reporting errors, do not for some cases. ER7318 "B COMPUTED WKS MISSED" is the sum of ER7264 "B61 # WKS OTR ILL (HD-E)", ER7269 "B64 # WKS SELF ILL(HD-E)", ER7274 "B67 # WK VACATION (HD-E)", ER7279 "B70 # WK ON STRIKE (H-E)", ER7284 "B73 # WK UNEMPLOYED(H-E)" and ER7301 "B76 #WK OUT LAB FRC(H-E)". ER7319 "B COMPUTED WKS WORKED" has a value of 52 minus the value of ER7318 "B COMPUTED WKS MISSED". Similar considerations apply in comparable variables in sections C, D and E.

The following variables have all missing data:

- * ER7127 (question A36), reason why the family neither owns nor rents the HU; and
- * ER8838 (question G111), a checkpoint for number of dependents.

Missing variables include:

- * employment status for individuals other than the Head or Wife/"Wife";
- * question G113, the number of persons dependent on this

family for more than half of their support.

2. Cross-Year Individual File Problem Variables

In the 1968-1996 early-release individual file the EMPLOYMENT STATUS variables, ER33011, ER33111, ER33211 and ER33311, for 1993, 1994, 1995 and 1996, respectively contain zeros for every person in the file. In addition ER33016, HAS MEDICAL COVERAGE? 93 and ER33017, HEALTH GOOD? 93 also contain zeros for every person in the file.

F. Comparability with 1992 Final-Release File and Earlier Final-Release Files

Beginning with the 1993 wave, the data were collected using CATI (Computer Assisted Telephone Interviewing). This meant that information about each question was collected electronically by the interviewer and, in effect, was coded at the time of data collection. Conversion to standardized units of measurement, formerly performed as part of our coding operation has not yet been done. As a result, the data in the early-release files much more directly resemble the answers to questionnaire questions than the 1992 and early years' data did. For example, instead of one variable indicating monthly rental expense, rent costs now exist as two variables: one for the dollar amount and one for the time unit, e.g., \$500 per month and \$100 per week are typical of responses to the question about rent payments. Therefore many of the 1993 through 1996 early-release variables are not directly equivalent to variables from 1992 and earlier waves.

As mentioned above, dollar amounts are often associated with time units in the early-release file. PLEASE BE AWARE THAT WE ARE NOT COMMITTED TO INCLUDING THESE COMPONENT AMOUNT AND TIME UNIT DATA AS PART OF THE MAIN FINAL-RELEASE FILE. Our current plans are to release final data that resemble as closely as possible our traditional data files. However, the amount-time unit (and similar) data collected in CATI but not generally part of our prior final files MAY, if it is not included in the final-release file, be available as a separate, subsidiary file so that analysts who desire this detail can access it.

Unlike data collected through 1992, the family data have NOT been cleaned with our manual economic edit process (nor have imputations been made), so you must convert these kinds of amounts into some sort of consistent unit for inter-case comparison and make decisions about handling missing data. In addition, we expect that values for quite a few cases will change when we do perform economic edit operations. For instance, time spent working, being laid off, unemployed, out of the labor force, etc., does not sum to 52 weeks per year in about 10 per cent of the cases in the early-release file! In addition, all time unit questions include an "other" code, as well as options for missing data; amounts associated with these "other" codes will be recoded from missing data or else imputed when the data are cleaned.

You can create cross-year files with variables needed for your particular analysis by merging the necessary information from the appropriate family files. The needed family identification numbers appear both in the final-release and early-release cross-year individual files. They also appear on the final-release family data files. While they do not appear on the early-release family data files, you can obtain them from the Head's record in the 1968-1996 early-release cross-year individual file (see Section G, step 1, above). Detailed instructions for the process of creating the cross-year files are included in the 1990, 1991 and 1992 family documentation volumes and are also available at our Internet site and are not repeated here.

Much of our usual inter-year consistency checking was performed for the early-release 1968-1996 cross-year individual file, so we expect the records in this file to remain relatively stable for final-release.

G. Additional Notes: Sample Supplements in 1993, 1994, 1995

We added a Latino sample of 2,043 families to the PSID in 1990. We continued to follow these families for five more years through 1995. They were not interviewed in 1996. This sample is described in detail in the 1990 documentation. It was derived from a sample selected and interviewed by Temple University Institute for Survey Research for their Latino National Political Survey (LNPS). The Latino addition was made congruent with our usual ID scheme and unique identifier formats. Latino sample cases are easily identified in the family and individual files by the code values for 1968 ID Number (V20302 in the 1992 family file, V30001 in the 1968-1992 individual file and ER30001 in the 1968-1996 early-release individual file) -- the Latino sample has code values in the range 7001-9043.

In 1992 several different kinds of recontacts were attempted. These are described in detail in the 1992 family documentation, but briefly, three groups were selected:

- * all 1991 nonresponse;
- * a random subset of SRC and Census sample members who had become nonresponse in 1990 or earlier; and
- * all of Temple University's Latino sample persons who were not successfully interviewed by us in 1990.

The successfully recontacted Latino families have 1968 ID Numbers in the range 9044-9308.

Our recontact effort for 1993 included the resurrection of many nonresponse sample persons who shared a 1968 ID number with families still responding in 1992, similar to the second group selected for 1992 as described above. But in contrast to this 1992 group, priority was given to families with connected individuals under age 18. All sample individuals within such a family were selected for recontact, even if they themselves were older.

The main focus of the 1994 recontact effort was to follow nonsample ex-spouses of sample members; these ex-spouses had one or more children with the sample members, and at least one of those children was expected to be under age 18 by 1994. In addition, recontacts were attempted with persons who had become nonresponse in 1992 or 1993, with nonresponse core sample persons who had no other family members still responding by 1993 (some of whom had become nonresponse as early as 1969), and with some children formerly designated nonsample but born to sample members since the study began. There were no recontact efforts in 1995 and 1996.

The 1993, 1994, 1995 and 1996 waves included a change in PSID following rules. We now follow all sample persons who leave home, regardless of age. So, for example, when a sample male Head leaves his nonsample wife and their sample children, we attempt an interview not only with him but also with her because her household contains their sample children. Beginning with the 1994 data collection, we also now consider as sample those children who are born to a sample parent in a year when the sample parent was not in an interviewed family.
