Notes on the 'Income Plus' Files

1994-1997 Family Income and Components Files

Yong-Seong Kim, Tecla Loup, Joseph Lupton, and Frank P. Stafford
March 11, 1999

FIRST REVISION April 2000
SECOND REVISION December 20, 2000

I. Introduction and Overview

These files contain data and documentation for the Panel Study of Income Dynamics' 1994, 1995, 1996, and 1997 Total Family Income and its components. These files also include the state of residence for each of these four waves as defined by the PSID and FIPS state codes. Originally, these files included the three-digit 1970 Census occupation and industry codes for head's and wife's/"wife's" current or most recent main jobs, for the 1996 wave only. In April 2000, these codes were added for the 1994 and 1995 waves as well. Coding for 1997 has not been completed as of this writing. We expect to post 1999 occupations with the 1999 files. Also in April 2000, both USDA and Census Needs Standards were added for all four waves. This file provides basic information about these 1994-1997 income files. We recommend that you carefully review the following material before proceeding with any work using the data files.

Over the last two years the PSID Staff have taken steps to develop new processing software to be applied to the (Computer Assisted Telephone Interviewing) CATI output for family income and its components. For those familiar with the complex PSID income sequences for head, wife and other family unit members, this will be recognized as an ambitious project. An initial version of the Income Processing Software (IPS) was completed in the early fall of 1997 (CATIINC 1.0), and a modified version was completed (CATIINC 2.0) in the late summer of 1998. A further revised version was developed (CATINC 3.0) in the summer of 2000. This CATI 3.0 is more sophisticated in identifying anomalous cases. For example, the newest IPS version incorporates information from prior waves (and, if available, subsequent waves) to identify potentially problematic cases of data entry error.

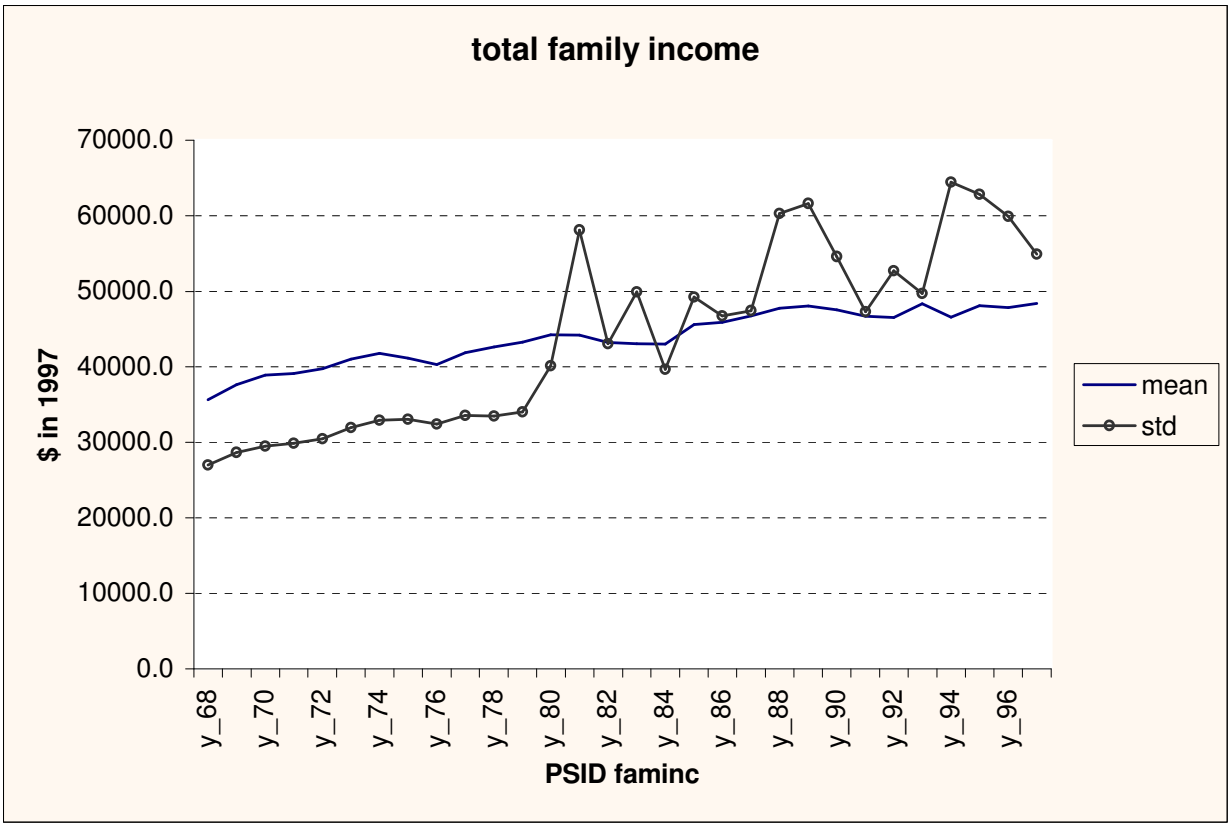
1. Early work with the income data from the PSID Income Plus files showed that the post-CATI data, 1993-1997, have a higher variance and that this seems to be concentrated in about 50 cases per year (out of 7,000 – 8,500 cases). Having extensively reworked our Income Processing Software (IPS) in the spring and summer of 2000 to create CATINC 3.0, we have recently applied it and edited additional cases. This new version of IPS is meant to work in tandem with our new editing system—more on this below. The IPS calculates income and its components in a straightforward way if all elements are there and there are no item non-responses or other anomalies in the underlying components. For the cases with unusual features one can calculate values from partial information. But this has limits. Also, simple keystroke errors in CATI can give rise to valid, but extreme values. To identify large, artificial changes in income

components—such as labor income—the program flags 'large changes,' particularly if not accompanied by a change in occupation and/or industry. These flagged cases are looked up to assess the larger context of the record including possible interviewer text fields of notes taken during the interview. If a better judgment edit can be made, a value is assigned. If not, a simple imputation may be used. Cases so modified are recorded as modified for the user.

2. The IPS output is intended to be then edited within an editing system. This system is described more fully below. In the interim we have effected a simpler (but time intensive) version of that system by creating Excel files of the input that goes into an income calculation and using that information for editing. This has lead to what we regard as very good income data for labor income and total family income. There are still some very minor improvements to the income files, 1994-1997 and 1999 which will be carried out for final archiving in our new edit system. How does our pre-archived income data look?

An overview of Total Family Income, 1968-1999 is provided in Figure 1 where we have carried out the application of the extensively revised IPS to process income from its detailed components and have carried out numerous checks for both cross-sectional and cross-wave outliers to identify potentially anomalous cases. IPS has also treated the SEO sample (see the PSID website under 'Overview' and then 'Sample') differently. In the case of imputations for missing data it has developed within-SEO sample conditional values to avoid imparting a systematic upward bias for those cases. Across all four waves, 1994-1997, there were about 800 cases deemed to need some case-by-case checking. (Sort of like challenges in the NFL.) Overall, about 200 cases were changed, 1994-1997, with more of the changes in 1994, and fewer per year going forward to 1997. These changes are mostly designed to improve the data quality for panel analysis, but show up as a modest reduction in the cross-sectional variance of income.

Figure 1



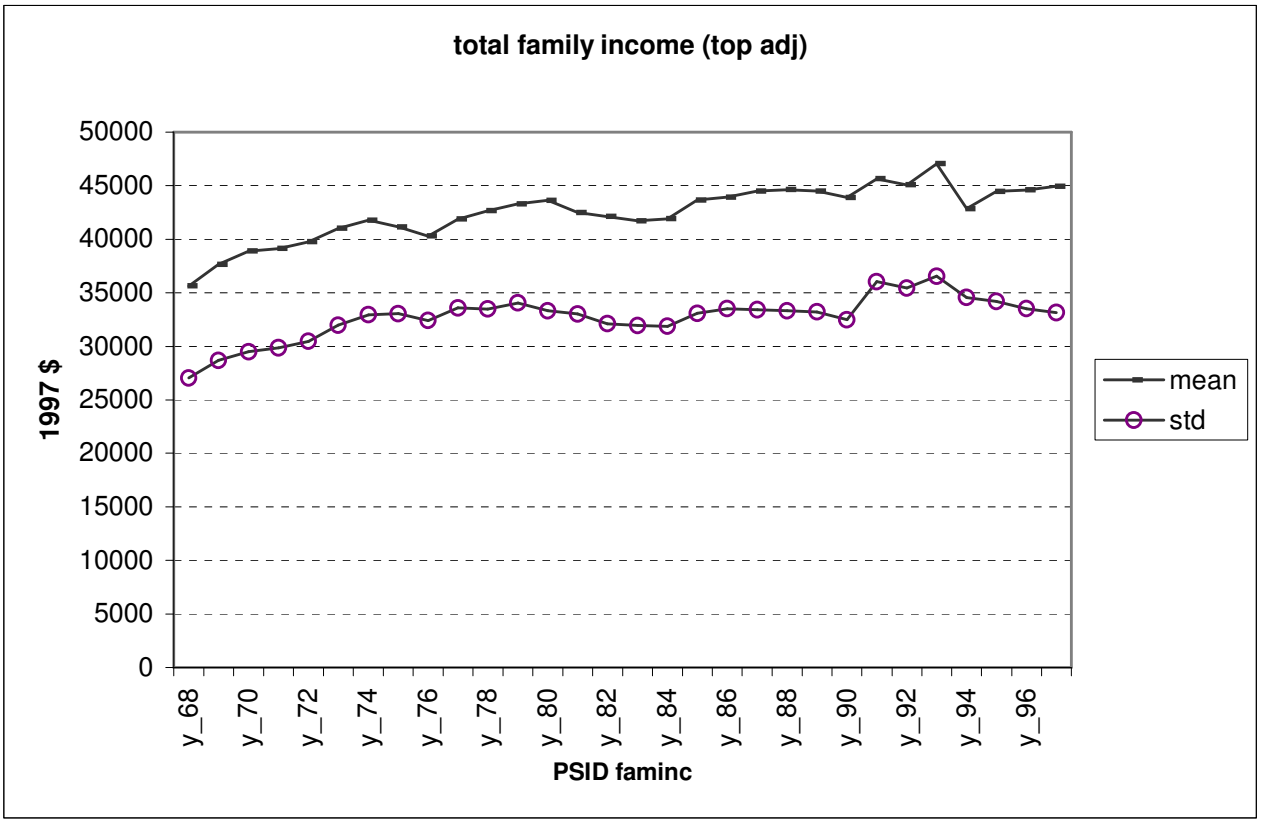
From Figure 1 there are two data 'eras': one 1968 to 1979 and then post 1980. Throughout there is a steady upward rise in (family weighted) total family income in 1997 CPI-U dollars, from $35,628 in 1967 (1968 survey year) to $48,394 in 1996 (1997 survey year). The standard deviation of family income appears smooth through time up to 1979 and then from 1980 forward is higher and more variable per year. The explanation for this change is that prior to 1979 the topcode value of income was $99,999. In 1980 it was increased to $999,999 and in 1981 it was increased to $9,999,999. The reasons for topcoding at $99,999 prior to 1980 appear to have been limits on logical record length in the early file structures of the PSID and a concern that top values may be dominated by reporting error. The current philosophy is to provide as many digits as needed to portray the full range of income, to check to be sure that insofar as possible these high values are valid and not keystroke errors, and allow the analyst to deal with issues of robust estimation.

For the period 1968 to 1993, income was also bottom coded at $1. This means that a few families with negative total family income, arising typically from a business or farm loss were set to $1. For 1994-1997 the Income Plus files allow a negative total family income. For the purpose of constructing Figure 1, the negative and 0 values for 1994 to 1997 were set to $1 to maintain comparability with pre-1994 bottom coding. One purpose for bottom coding appears to be simplicity in calculation income-to-needs ratios, a key goal in the study, particularly in its early years when the central focus was the study of poverty and income needs of the U.S. population.

The data in Figure 1 show a higher standard deviation after the top coding change in 1979-1980. This is not surprising, and Figure 2 shows the total family income series 1968-1997 if the

topcoding at $99,999 had been extended throughout. In that case the standard deviation is always below the mean and exhibits relatively little year-to-year variation throughout. The standard deviation in total family income in the Income Plus files for 1994-1997 seems quite in line with the long term pattern.

Figure 2



A concern is that even with careful data processing with IPS and individual case editing, the CATI based income could be dramatically different from paper and pencil and individual editing prior to 1993. One simple check is suggested by referring back to Figure 1. Throughout the period when topcoding was extended to $9,999,999 there are much larger year-to-year variations in the cross section standard deviation of total family income. Further there are three post-1979 'episodes' of notably higher standard deviation values: the early 1980's recession and recovery of 1981-1983, the late 1980's expansion of 1987-1989, and the post-Gulf War expansion of 1994-1996. There does not appear to be a pure CATI effect, since 1993 differs from 1994-1996 and 1997 is about in the same range as 1993. Within the CATI years, then 1993 and 1997 are 'low' in terms of post-1980 values for the standard deviation, and 1994-1996 are 'high' in these same terms.

Despite the apparent similarity in the overall cross sectional income measures (after an enormous effort to maintain comparability in the face of extensive technical changes) the output from each year since the 1992 survey has had its own set of special input problems, owing to successive changes and improvements in the CATI software. In 1993 there was a migration from the

traditional paper questionnaire to an initial CATI application. This has the potential to have created a seam problem in the income data, given the complex structure of the PSID question sequences designed to measure income at the detailed component level by person within the family. In processing the 1993 family income there was a need to create an initial version of our software system which paralleled the pre-1993 system of software and editing of paper questionnaires. This is necessarily different in numerous details from the prior system, again creating a potential seam in the income data. For 1994 there were additional changes to CATI and the 1994 data in these files, as noted above, have been processed by the CATIINC 3.0 software. From 1995-1997 the CATI system had been stabilized and we are applying the same software for these files, so these 1994/5-1997 family income data should be quite consistent with one another.

The issue for researchers to consider then is a potential double seam created by a shift to CATI and a concomitant change in the processing software in 1993 and 1994. It is our judgment that the seam problem created by CATI collection itself was quite minor. To illustrate, a 1994 study of recording errors in 200 cases assigned to CATI and 200 cases in paper and pencil shows similar rates of recording errors. The potentially far more important source of non-comparability is in the processing system – the system which takes partial and otherwise anomalous cases and reaches an estimate or imputation for a 'final value'.

Taking the processed 1994-1997 output from CATINC 3.0, cases were reviewed and adjusted, if that was deemed best. For example, by misreporting the `per' variable, a respondent could have $75,000 `per year' miscoded as $75,000 `per week'. A file of these corrections was kept, by year, so that those wishing to apply alternative editing judgments can do so. In the paragraphs below we provide the user with a selection of the extreme cases which we `hand edited' after having employed our program to compute family income and its components.

We have recently (September - December of 2000) reviewed the labor income of head and wife and have `hand edited' an additional set of the cases with high values of transfer income. The components of transfer income are very diverse. This has made it difficult to carefully evaluate outliers without a case by case review of the interview and interviewer thumbnail comments—plus information on the family from prior and subsequent years. Analysts are urged to exclude or otherwise allow for cases in which the value of transfer income is above the top 99.5 percentile point. Also, analysts may want to either exclude or allow for these cases n the analysis of income totals which have such extreme values in them (i.e. total family income, taxable income of the head and wife) or to assign the transfer income part of these family aggregates to zero or the median of transfer income.

ILLUSTRATIVE POST SOFTWARE CASE `HAND' OR JUDGMENTAL EDITING:

1997 ID=02314 had Social Security income of $1,214,952. The head reported $15,000 a year, and the wife reported $99,996 per month! Checking the marginal notes, we read that the COMBINED head and wife Social Security income was $15,000 for the year; the interviewer entered this amount for the wife intending to alert us to the fact that head and wife couldn't separate their respective amounts. (It worked, but this is not what she would have been advised

to do—yet she made this judgment on the spot to continue the interview.) We recoded the amount to $15,000.

1997 ID=01368 had head and wife transfer income of $347,041. This was due to an `other retirement pension' amount of $308,712. From other evidence, the amount seemed unlikely for this family. We checked back to 1996, where the same item was reported as $3,087. It appears that the 1997 other pension reported amount is missing a decimal and should be $3087.12, or rounded to $3087. In addition, this family reported $38,329 from "anything else" (miscellaneous transfers), with no marginal note explaining the source. Checking the 1996 and 1995 interviews, no similar amount had been reported. We judged that the $38,329 is a lump sum payment of some kind, probably health insurance reimbursements—the head has been ill, in and out of nursing homes and hospitals. So their total transfers were entered as $3087 for 1997.

1997 ID=05135 had other family unit member (OFUM) transfer income of $221,132. Two OFUMs, a son and a girlfriend, were in this family unit (FU). The son reported $511/month of Supplemental Security Income (SSI), received all year, for a total of $6132. The girlfriend reported "anything else" of $10,000/week received in Jan, Mar, June, August, and November. Aside from the fact that it appears she means $10,000 per year, this money is from student loans and should be entirely removed from income. We never consider loans as income, period. So total OFUM transfers for this case=$6,132.

1996 ID=00427 had head and wife transfer income of $238,446. Here are all the non-labor components: the `wife' reported $500/yr dividends—which seems ok, $2000/yr interest—which seems ok, but $18,000 per WEEK help from relatives! This is not ok! Checking back to 1995, she reported $20,000 per year for the same thing. So the `per' time unit was judged to be `per year' rather than `per week'. In addition, head reported $18,000 per year in help from non-relatives. The marginal notes make it clear that this is double-counted with the `wife's'. So their total transfers were judged to be $20,500 per year.

Cases of this sort were quite rare. When it is remembered that the field-based input on the part on interviewers for these files consisted of approximately 1.5 million data entries, the presence of some outlier cases is not surprising. Interviewing is a delicate art of continuing to engage a respondent who is providing a lengthy and complex interview via telephone. Some errors are inevitable, and not all will be caught in our processing and hand editing systems. Users need to keep this in mind and are advised to calculate descriptive statistics and quantiles on the groups and subgroups in their analyses. Users should also realize that in looking at individual cases, changes were made only if there were fairly obvious indications of misreporting (as illustrated above). Our approach is not nor has it ever been to second guess respondent reports or stylize the economic reality of the families in our study.

The tradition in the PSID has been to set income data with negative or zero total family income values to $1. For these files of the 1994-1997 years, those with negative or zero total family income (arising, e.g., from a business loss or from living on liquidated assets) were simply left as negative or zero values. As discussed above users wishing to increase time series consistency may want to allow for this difference.

While we have listed and adjusted obvious outliers, there are two remaining data issues which analysts should recognize.

There are 519 working wives who were skipped past the labor force CATI checkpoint in the income question series in the 1994 Survey. In the process of revising the Income Plus file, the income variable for those respondents was first flagged and was imputed using various methods including cross-year imputation. As a result, the 1994 cases of wife earnings that were treated by the IPS is increased relative to other years (see Section V: Income Editing Procedures).

We have made numerous data checks by looking at time series of descriptive statistics for the years 1968-1997. Nonetheless, this is a new and complex system and income is itself highly dispersed. For analysts, robust estimation and descriptive techniques should be the order of the day.

Finally, we note that the data used in this paper are all in 1997 dollars (CPI-U, 1982-1984 = 100). (The data in the released files are *not* in 1997 dollars, rather they are in actual amounts). The relevant CPI-U values are:

1997: 160.5
1996: 156.9
1995: 152.4
1994: 148.2
1993: 144.5
1992: 140.3
1991: 136.2

The conversion factors which were applied to the data by (calendar) year data were:

1997: 160.5/156.9 = 1.02295
1996: 160.5/152.4 = 1.05315
1995: 160.5/148.2 = 1.08300
1994: 160.5/144.5 = 1.10727

For those working with data from 1991-1993 to study time series patterns the conversion factors would be:

1993: 160.5/140.3 = 1.14398
1992: 160.5/136.2 = 1.17841

There was a study comparing the PSID and CPS, 1969-1991, by Greg Duncan, Timothy Smeeding and Willard Rodgers (Household Income Dynamics in the 1970's and 1980's, working paper, April 24, 1995), based on families headed by a 25-54 year old. The study showed that, particularly for the Gulf War Recession year, 1991, the PSID and CPS do line up relative to their values as of 1969. However, in the mid to late 1980's the PSID (with 1969 as the reference year) was showing substantially lower income at the lower (20 and lower) percentile points. One interpretation is that while the PSID gets a more complete income picture (more components and

detail and therefore less underreporting), in more recent years we also get a better response from people with low income and assets. This has come up in the wealth area where, in contrast to evidence from some other studies, we show that African Americans are less likely to have a bank account in the 1990's compared to the 1980's (see Erik Hurst, Ming Ching Luoh and Frank Stafford, "Wealth Dynamics of American Families, 1984-1994," Brookings Papers on Economic Activity," (1998:I), p. 296-301). These low income and low asset ownership respondents, one can hypothesize, actually report how low their income and assets are in the PSID.

In summary, there are several points to bear in mind in CPS/PSID household/family income comparisons. 1. 1991 was a special (vintage) year in terms of CPS/ PSID alignment. 2. Both studies have potentially significant data seam problems in the 1990's. The post 1992 survey year CATI and processing of CATI issues for the PSID have been noted above. 3. At the median there appears to be an approximately a 0-5% PSID/CPS difference in medians between CPS and PSID for 1994(93) through 1997(96).

Table 1
Median Family Income PSID and CPS Comparison

	PSID Median (1997$)		**CPS Median (1997$) ***
	Release of		
	March 1999	**December 2000**	
1993 (1994 Survey)	$32,749	$33,321	$34,432
1994	$34,607	$34,654	$34,807
1995	$35,070	$35,173	$35,807
1996 ****	$34,985	$34,903	$35,979

* Source: http://www.census.gov/hhes/income/dinctabs.html

** In 1997, we were forced to drop 2,843 of the 3,967 families in the SEO (low income) sample because we were not successful in our efforts to secure funding for the continued data collection for this part of the core sample. The sample suspension as of 1997 and every other year interviewing changes after 1997 are discussed in our April, 1997 PSID Newsletter and our April, 1998 PSID Newsletter (viewable at this website). The 1997 family weights have been designed to apply to either the remaining pre-1997 PSID core (allowing for the 2,843 dropped SEO cases) or the 1997 core, which now includes the new sample of post-1968 immigrant (P68I) families.

II. Data

There are four data files: the 1994, 1995, 1996 and 1997 Family Income and Components Files. Each file contains information about Total Family Income, which is the sum of Taxable Income of the Head and Wife plus Transfer Income of the Head and Wife plus Taxable Income of Other Family Unit Members plus Transfer Income of Other Family Unit Members plus Social Security Income.

Also included are: Farm Income, Business Income, Labor Income of the Head, Labor Income of the Wife, and (for 1994) whether the labor income of the wife was assigned by imputation (as the

product of 1994 Survey Year hours times 1993 Survey Year wage rate). The state of residence, as defined by the PSID and FIPS state codes, is also included for each wave. Occupation and industry codes are added to the 1994-1996 files; these variables were not included as part of those early-release main files because they had not yet been coded when those files were released. In addition, the revised income plus file includes various components of labor income so that it helps users to identify which part of labor income might be imputed. Also variables for number of businesses owned by the family, head, and wife were generated as a by-product of the IPS. Finally, not only the total business income but also the disaggregation of labor and assets parts of the business income for Head and Wife/"Wife" are generated.

Finally, both USDA and Census needs standards have been added for all four years. Note that a very few cases have missing data in these variables (1 case in 1994, 3 in 1995, 6 in 1996, and 1 in 1997). The early release family data files, which match the Income Plus files, included some family records that will not appear on the archive releases of the family files. (Cases can be invalidated for various reasons, e.g., a splitoff interview was found to be invalid because the splitoff had moved back home before the splitoff interview was taken.) The needs variables, however, were calculated using the cross-year individual file, which did not include individuals from those invalidated families.

These 1994-1997 Family Income Files contain one record for each family interviewed in 1994-1997. For each year, the file includes a special sample of recontacted respondents (notably numerous in 1994), as part of a large methodology study. These 1994-1997 files do not include the special Latino sample. The case count of families in the 1994 Family Income File is 8659. For 1994 the case count of families that have a non-zero family panel weight (see the weights files for 1994-1996 released 9/98) was 7747. The difference is the consequence of the recontact families. They can be used for some analysis purposes, but simply have a zero family weight. Parallel differences of this sort exist for 1995 and 1996. Users wishing to apply FAMILY WEIGHTS in their analysis will need to visit the weight section of the data library (PSID Data Files, 1993-1997). (We do not place the Family weights here for the sake of controlling redundancy in our data library and website.)

The 1997 weights are complicated by sample suspension and the addition of a refresher sample of post-1968 immigrants, but they are now available and can be applied to these 1997 family income variables.

These 1994-1997 data may be subject to relatively minor changes once the archive-release versions of the 1994, 1995, 1996 and 1997 family files become available. The data are in raw ASCII form. Refer to the data definition statements—SAS or SPSS—for record format layout information, variable names, variable labels, and missing data codes.

Table 2
File Attributes and Variables for Data Files

File Name	Records	LRECL	Num of Variables
FAMINC94	8,659	229	48
FAMINC95	8,570	228	48
FAMINC96	8,517	225	48

FAMINC97	6,307	215	44

III. SAS & SPSS Data Definition Statements

These files contain SAS and SPSS data definition statements providing information about the variables in the data files. Two files, one of each type, SAS and SPSS, corresponding to each data file, are provided. The naming conventions are the same as for the data files, e.g., FAMINC94.SAS contains SAS statements for the 1994 Family Income data file, and FAMINC94.SPS contains SPSS statements for the 1994 Family Income data file. Similar files also exist for 1995-1997.

The data definition statements provide variable names, variable labels, locations. These processed files have no `missing' data.

The SAS and SPSS data definition 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 SAS or SPSS statements needed to complete a program. Users wishing to migrate to other formats may use a commercial software for such purposes, such as STAT/TRANSFER.

IV. Documentation

The following machine-readable documentation files are provided for the designated data files.

Table 3
File Description and Attributes for Documentation Files

Contents	Filename	Approximate Printed pages
Notes on the 'Income Plus' Files	YPlsNote.doc	16

V. Income Editing Procedures

General Editing Rules for Labor Income

In a given year of PSID, the data were first screened to detect potentially problematic values for each component of the head and wife's labor income variables. Programs were written to flag: 1) Don't Know (DK) and Not Applicable (NA) (e.g., 9999998, 9999999) ; 2) 'short' or less than full field width codes (e.g. 9999, 9998); and 3) keystroke errors (e.g. 9996, 88889). The resulting numbers of such instances for the head and wife are reported in Tables 4 and 5.

Table 4
Number of flagged cases in head's labor income

PSID	FRM	WG	BNS	OVT	TIPS	COM	UK	PRA	GAD	RM	XTR	Total
94	30	579	33	81	9	12	0	19	5	0	29	797
95	30	454	20	65	11	11	0	11	1	1	13	617
96	18	381	27	47	5	6	0	8	4	1	16	513
97	20	285	31	33	6	6	6	3	3	1	8	402
Total	98	1699	111	226	31	35	6	41	13	3	66	2329

* FRM = Farm income, WG=Wage and salaries, BNS=Bonus, OVT=Overtime, COM=commissions, UK=Unknown Jobs income, PRA=Professional practice, GAD=Market Gardening, RM=Roomers and Boarders, XTR=Extra job income.

Table 5
Number of flagged cases in wife's wage income

PSID	WG
94	732
95	160
96	156
97	145
Total	1193

Needless to say, these values must be replaced by real values so that any level of aggregation of labor income components yields meaningful amounts. To obtain an informed value, the following data elements were examined: 1) cross-sectional information such as from the employment section; 2) cross-year information for the same family; 3) the cross-sectional distribution from the current year employment section. When all these efforts failed to deliver a candidate value, interviewers' notes were read for a possible explanation.

As indicated in Tables 4 and 5, the two most problematic components are "Wages and Salaries" and "Overtime". The procedures for these two components are discussed below.

1. Wages and Salaries:

Wages and salaries reported in G13 and G52 should not include income from an unincorporated business. For such cases a zero amount was assigned and the amount was relocated to the business income section. There is useful contextual information for wages and salaries (G13 and G52) in the employment sections (B, C, D and E). By definition, wage and salary income must be closely related to hours of work.

In the case where G13 or G52 had a potential problem, the characteristics of all jobs (excluding unincorporated business jobs) reported by the head (wife) were examined. For each job, wage and salary were calculated based on hours of work (= how many weeks work times how many hours per week) and the hourly wage rate. When the number of weeks was not reported, the month strings of the job (weeks = # of months worked times 4.33) were used as a proxy. Then, wage and salary income of each job that head had reported was aggregated. In some cases, hours of work and wage and salaries of other years were compared. In many cases, the wage rate (after CPI adjustment) and hours of work in other years provide useful information in imputing amounts. The variables from the employment sections, month strings for work history, were a central input to editing.

For the remaining cases, reference was made to the cross-sectional distribution of the variable. For sake of continuity, the assignment method documented in previous years of the PSID was generally followed. The basic principle is that cases are grouped jointly by several categories such as occupation, home ownership, and type of housing. In this revision of income processing, geographical characteristics (urban vs. non-urban) were also added in the assignment. Figures 3 and 4 show the assignment groups. For each group, the mean of hourly wage rates was calculated (as reported in Tables 6 and 7). To impute amounts for missing cases, hourly wage rates from corresponding groups were used.

Finally, when all the above methods failed, the median of wage and salaries was substituted. The number of cases for median substitution are: 677 in 1994, 202 in 1995, 227 in 1996, and 0 in 1997.

2. Overtime:

In the cases where overtime amounts needed to be replaced, first wage rate and hours of work in Sections B and D were examined. If the head and wife were currently employed, respondents are asked questions about wage rates for any extra hours. For example, in the case where the head's main job is salaried or commissioned, B15 asks "About how much would you make per hour for those extra hours?" (with choices of 1 = time and half, 2 = double time, …). A similar question appears in B17 when the head's main job is not salaried or commissioned. Along with extra work hours and hourly wage rate, this information can be used to impute a value of overtime earnings.

In some cases, the reported hours for overtime work (B81) was an excessively large number. Probably, some respondents were confused between overtime hours and excessive hours. To deal with such cases, a restriction was imposed such that overtime work = min {10% of total work hours, overtime hours reported}. For cross-year comparisons, ratio of overtime earnings to earnings in other years was calculated, and this ratio was used in the imputation.

Figure 3
Assignment Groups: Hourly wage rate for Head:

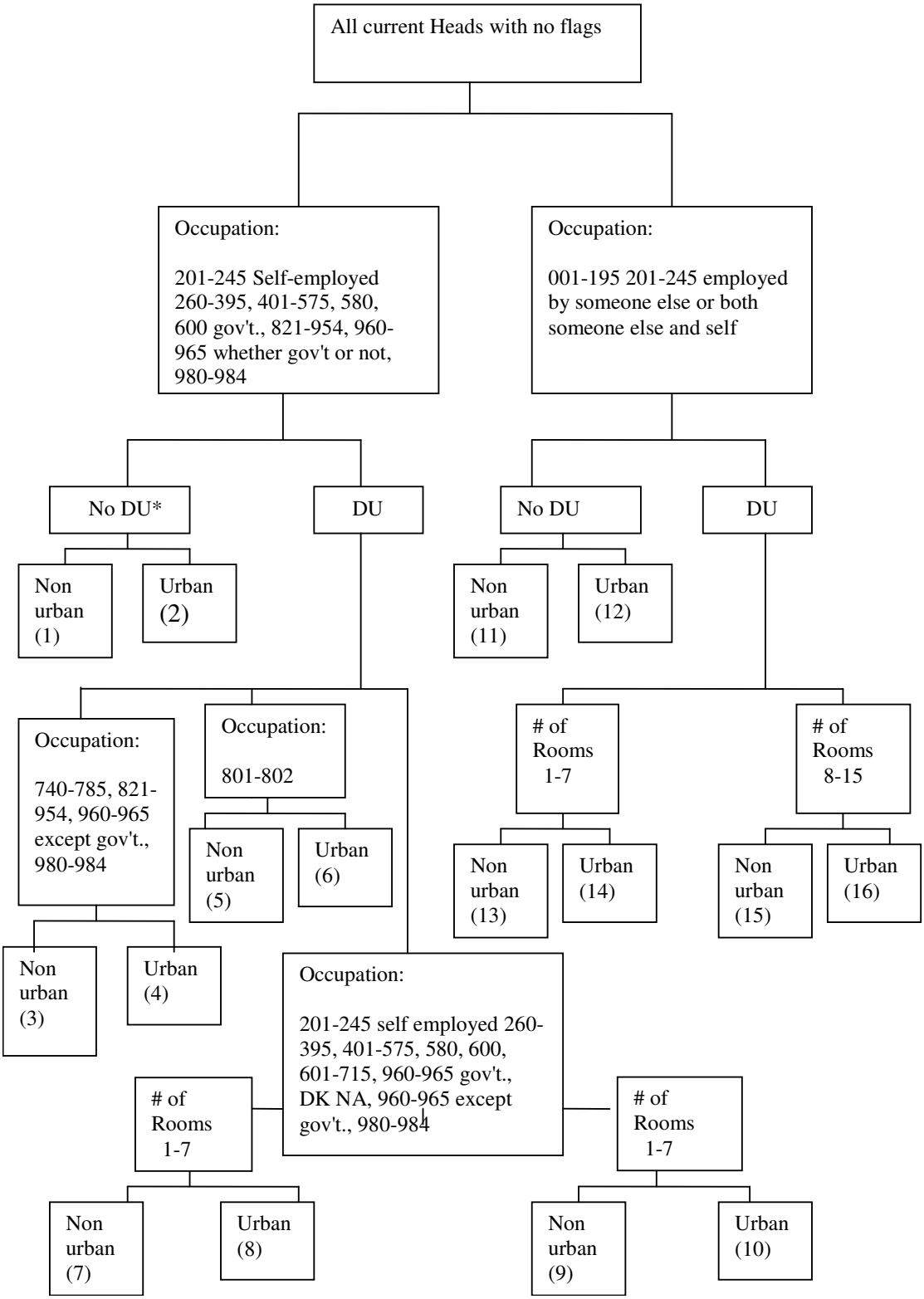

All current Heads with no flags
Occupation:
201-245 Self-employed 260-395, 401-575, 580, 600 gov't., 821-954, 960-965 whether gov't or not, 980-984
Occupation:
001-195 201-245 employed by someone else or both someone else and self
No DU*
DU
No DU
DU
Non urban (1)
Urban (2)
Non urban (11)
Urban (12)
Occupation:
740-785, 821-954, 960-965 except gov't., 980-984
Occupation:
801-802
# of Rooms 1-7
# of Rooms 8-15
Non urban (5)
Urban (6)
Non urban (13)
Urban (14)
Non urban (15)
Urban (16)
Non urban (3)
Urban (4)
Occupation:
201-245 self employed 260-395, 401-575, 580, 600, 601-715, 960-965 gov't., DK NA, 960-965 except gov't., 980-984
# of Rooms 1-7
# of Rooms 1-7
Non urban (7)
Urban (8)
Non urban (9)
Urban (10)

Table 6
Assignments for Head's Hourly Wage Rate

PSID 94

group	n	Min	max	mean	std
1	371	1.07	147.66	8.47	10.00
2	2081	1.00	162.67	9.60	9.35
3	120	1.35	166.88	9.19	15.38
4	324	1.00	113.75	10.76	10.80
7	411	1.29	100.00	11.60	8.92
8	1261	1.37	161.85	14.48	10.22
9	84	1.09	50.60	14.41	8.58
10	320	2.08	183.33	20.48	18.57
11	65	1.52	63.49	11.64	8.91
12	525	1.32	115.61	14.42	10.31
13	127	2.18	48.25	15.01	8.19
14	644	1.63	176.00	21.08	16.24
15	58	1.16	52.00	17.68	10.43
16	409	3.40	178.57	30.13	23.09

PSID 95

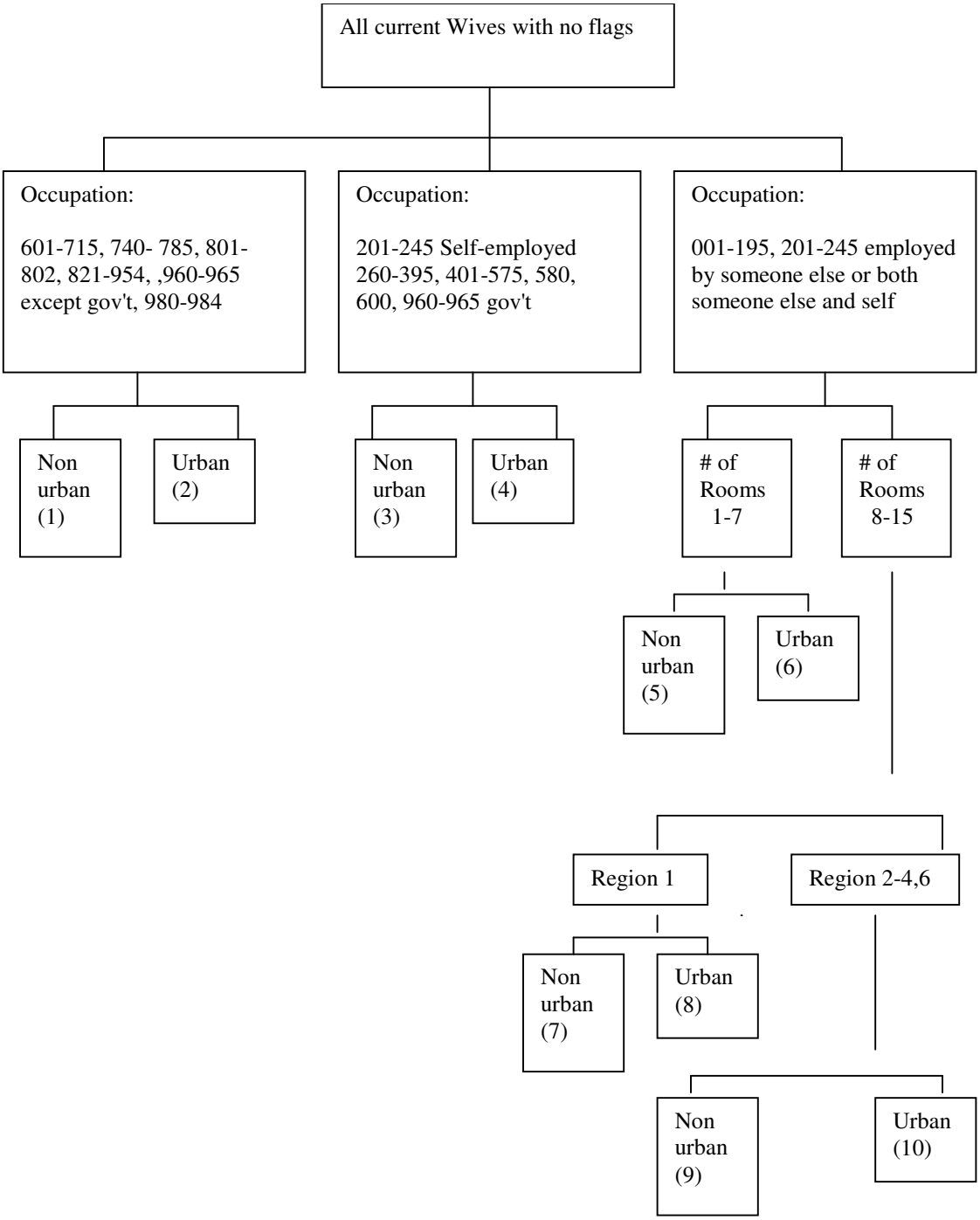
group	n	Min	max	mean	std
1	422	1.07	94.57	9.22	7.73
2	2012	1.02	104.25	9.46	8.20
3	69	1.69	42.64	9.06	6.33
4	377	1.21	91.43	10.38	10.04
7	299	1.65	125.00	13.82	9.90
8	1324	1.03	182.55	14.38	11.80
9	94	1.25	171.22	21.92	19.66
10	338	2.64	146.94	19.29	14.52
11	111	1.69	83.56	14.76	12.28
12	463	1.02	120.45	14.70	11.26
13	165	1.04	90.91	18.97	10.64
15	96	3.85	170.07	29.44	24.35
16	345	2.92	191.49	28.40	19.88

PSID 97

group	n	Min	max	mean	std
1	342	1.12	53.12	8.21	6.10
2	1612	1.00	96.34	9.85	7.95
3	114	1.07	86.09	9.56	8.66
4	229	1.11	44.80	10.94	6.16
7	377	1.01	105.47	12.47	9.27
8	955	1.15	192.46	15.60	12.42
9	85	1.44	37.12	13.95	7.36
10	305	2.50	163.27	22.94	20.12
11	79	1.50	57.33	11.17	8.23
12	471	1.56	175.53	14.47	11.47
13	125	1.46	35.42	14.40	6.10
14	556	2.05	150.00	20.84	13.42
15	46	2.61	64.00	21.55	12.31

16	384	2.25	191.49	30.55	22.18

Figure 4
Assignment Groups: Hourly Wage Rate for Wives:



Region 1 = Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont

Region 2 = Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, Wisconsin
Region 3 = Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, Washington, D.C., West Virginia
Region 4 = Arizona, California, Colorado, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, Wyoming
Region 5 = Alaska, Hawaii
Region 6 = others

Table 7
Assignments for Wives Hourly Wage Rates

PSID94

group	n	min	max	rate	std
1	173	1.03	5.49	2.45	1.04
2	477	1.00	14.27	2.70	1.64
3	162	1.03	11.90	2.99	1.48
4	856	1.03	40.35	3.55	2.58
5	60	1.03	9.79	3.41	1.87
6	336	1.01	14.59	4.72	2.51
7	60	1.03	9.10	4.39	1.87
8	271	1.03	42.93	5.98	3.79
9	6	5.24	8.93	6.88	1.52
10	80	1.03	13.56	6.36	3.30

PSID95

group	n	min	max	rate	std
1	122	1.03	5.75	2.28	1.01
2	621	1.01	90.23	2.82	3.79
3	233	1.01	12.88	3.45	2.07
4	931	1.03	37.77	3.51	2.37
5	101	1.12	15.43	4.30	2.58
6	370	1.03	30.19	4.67	2.94
7	111	1.03	14.59	5.47	2.78
8	327	1.03	27.47	5.86	3.33
9	8	3.43	9.44	6.86	1.90
10	88	1.16	24.04	5.48	3.11

PSID96

group	n	min	max	rate	std
1	195	1.03	12.02	2.58	1.24
2	361	1.01	9.44	2.94	1.63
3	195	1.03	10.00	2.98	1.53
4	801	1.03	38.63	3.86	2.63
5	87	1.03	9.39	3.81	1.95
6	359	1.03	16.23	4.78	2.47
7	86	1.03	13.05	4.15	2.26
8	328	1.03	24.04	5.92	3.29
9	9	3.92	10.04	7.28	2.12
10	89	1.03	47.22	6.63	5.51