

1996 DIARY SURVEY
PUBLIC USE MICRODATA
DOCUMENTATION

June 15, 1999

TABLE OF CONTENTS

I.INTRODUCTION.....	4
II.CHANGES FROM THE 1995 MICRODATA FILES	4
III.FILE INFORMATION	7
A.DATA SET NAMES	8
B.RECORD COUNTS PER QUARTER	9
C.DATA FLAGS.....	9
D.FILE NOTATION	10
E. DETAILED VARIABLE DESCRIPTIONS.....	11
1. CONSUMER UNIT CHARACTERISTICS AND INCOME FILE (FMLY).....	11
a. CU AND DIARY IDENTIFIERS	11
b. CU CHARACTERISTICS.....	12
c. CHARACTERISTICS OF REFERENCE PERSON AND SPOUSE	16
d. WORK EXPERIENCE OF REFERENCE PERSON AND SPOUSE	19
e. INCOME.....	21
f. OTHER MONEY RECEIPTS	24
g. TAXES.....	25
h. RETIREMENT AND PENSION DEDUCTIONS	27
i. FOOD STAMPS.....	28
j. FREE MEALS AND GROCERIES	30
k. HOUSING STRUCTURE	31
l. WEIGHTS	32
m. SUMMARY EXPENDITURE DATA	34
2. MEMBER CHARACTERISTICS AND INCOME FILE (MEMB)	35
a. CU AND MEMBER IDENTIFIERS	36
b. CHARACTERISTICS OF MEMBERS	36
c. WORK EXPERIENCE OF MEMBERS.....	39
d. INCOME.....	40
e. TAXES.....	42
f. RETIREMENT AND PENSION DEDUCTIONS.....	43
3. DETAILED EXPENDITURES (EXPN) FILE	44
4. INCOME (DTAB) FILE.....	45
5. PROCESSING FILES	45
a. AGGregation file.....	45
b. LABEL file.....	46
c. UCC file.....	46
d. SAMPLE program file	46
IV. TOPCODING AND OTHER NONDISCLOSURE REQUIREMENTS	47
A. CU CHARACTERISTICS AND INCOME FILE (FMLY)	47
B. MEMBER CHARACTERISTICS AND INCOME FILE (MEMB).....	51
C. DETAILED EXPENDITURE FILE (EXPN).....	52
D. INCOME FILE (DTAB)	53
V. ESTIMATION PROCEDURES.....	54
A. DEFINITION OF TERMS	55
B. ESTIMATION OF TOTAL AND MEAN EXPENDITURES.....	55
C. ESTIMATION OF MEAN ANNUAL INCOME	57
VI. RELIABILITY STATEMENT	57
A. DESCRIPTION OF SAMPLING ERROR AND NONSAMPLING ERROR	57
B. ESTIMATING SAMPLING ERROR.....	58
1. VARIANCE ESTIMATION.....	58
2. STANDARD ERROR OF THE MEAN	59
3. STANDARD ERROR OF THE DIFFERENCE BETWEEN TWO MEANS	59
VII. MICRODATA VERIFICATION AND ESTIMATION METHODOLOGY	61

A. SAMPLE PROGRAM	62
B. OUTPUT	76
VIII. DESCRIPTION OF THE SURVEY	78
IX. DATA COLLECTION AND PROCESSING.....	78
A. BUREAU OF THE CENSUS ACTIVITIES	79
B. BUREAU OF LABOR STATISTICS ACTIVITIES	79
X. SAMPLING STATEMENT.....	79
A. SURVEY SAMPLE DESIGN.....	79
B. COOPERATION LEVELS.....	80
C. WEIGHTING.....	81
D. STATE IDENTIFIER.....	81
XI. INTERPRETING THE DATA.....	82
XII. APPENDIX 1--GLOSSARY.....	82
XIII. APPENDIX 2 -- UNIVERSAL CLASSIFICATION CODE (UCC) TITLES.....	83
A. EXPENDITURE UCC's ON EXPN FILE.....	83
B. INCOME AND RELATED UCC's ON DTAB FILE.....	93
XIV. APPENDIX 3 -- UCC AGGREGATION	95
XV. APPENDIX 4 -- FMLY AND MEMB VARIABLES ORDERED BY START POSITION.....	96
A. FMLY FILE	96
B. MEMB FILE.....	98
XVI. APPENDIX 5 -- PUBLICATIONS AND DATA RELEASES	99
XVII. INQUIRIES, SUGGESTIONS, AND COMMENTS.....	101

I. INTRODUCTION

The Consumer Expenditure Survey (CE) program provides a continuous and comprehensive flow of data on the buying habits of American consumers. These data are used widely in economic research and analysis, and in support of revisions of the Consumer Price Index. To meet the needs of users, the Bureau of Labor Statistics (BLS) produces population estimates (for consumer units) of average expenditures in news releases, reports, bulletins, articles in the Monthly Labor Review, and on diskettes. Tabulated CE data are also available on the Internet and by facsimile transmission (see appendix 5). The microdata are available on public-use computer tapes (pre-1996) or compact disk-ROM (CD-ROM).

The Diary microdata files present detailed income and expenditure data for the Diary component of the CE for 1996. Beginning with the 1996 release, SAS data sets, as well as ASCII files, will be made available on CD-ROM. Also beginning with the 1996 release, CE microdata will no longer be available on magnetic tape. Estimates of average expenditures from the Diary survey, integrated with data from the Interview survey, are published in *Consumer Expenditures in 1996, Report 926 (1998)*. A list of recent publications containing data from the CE appears at the end of this documentation.

The microdata files are in the public domain and with appropriate credit, may be reproduced without permission. A suggested citation is: "U.S. Department of Labor, Bureau of Labor Statistics, Consumer Expenditure Survey, Diary Survey, 1996".

II. CHANGES FROM THE 1995 MICRODATA FILES

Several major changes have taken place from the 1995 release. Variables whose content includes "year" information have increased in length to achieve Y2K compliance. Since many start positions changed because of this, we have also taken the opportunity to eliminate empty spaces in the data files which have accumulated over the years. Please be aware that many variables have different positions in the 1996 data files than they did in previous years.

There have also been major revisions to the topcoding procedures. Please refer to section IV for information about the new topcoding methodology and for a comprehensive list of changes and affected variables.

There was a sample redesign in 1996. The sampling frame is now generated from the 1990 Census of Population 100-percent-detail file.

Finally, the CU weighting procedure has been slightly modified. The new procedure is outlined in section X.C.

Other changes from the 1995 microdata files follow.

- 1) A new topcoding methodology is in place with the 1996 microdata release. See section IV for details on the new methodology and new topcode values. Major topcoding changes are as follows:
 - REGION and POPSIZE are no longer subject to suppression.
 - STATE will include some "re-coded" states. These are observations for which the state code is replaced by the code of another state.
 - STATE_, a flag variable for STATE has been created. It can have the following values.
 - 'D' -- STATE contains an unaltered code.
 - 'T' -- STATE is suppressed (blanked) out due to non-disclosure requirements.
 - 'R' -- 1) STATE has been re-coded for that observation or 2) that state contains some re-coded observations from other states.

- 2) The following variable has been deleted from the FMLY files.

BASEWTA The inverse probability of selection for the CU adjusted for subsampling in the
field -- BLS derived.

- 3) The following variables have been added to the FMLY files.

POVERTY Is CU income below current year's poverty threshold?
POVERTY_ POVERTY flag

- 4) The following variables in the FMLY files have code and code definition changes.

The following changes apply to EDUC_REF and EDUCA2:

The codes eliminated are:

- 1 Elementary (1-8 years)
- 2 High school, less than H.S. graduate
- 3 High school graduate
- 4 College, less than College graduate
- 5 College graduate
- 6 Graduate school
- 7 Never attended school

The new codes that apply are:

- 00 Never attended school
- 10 First through eighth grade
- 11 Ninth through twelve grade (no H.S. diploma)
- 12 High school graduate
- 13 College, less than college graduate
- 14 AA degree (occupational/vocational or academic)
- 15 Bachelors degree
- 16 Masters degree
- 17 Professional/doctorate degree

The following changes apply to DESCRIP:

Code 10 (*Unoccupied site for mobile home, trailer or tent*) has been changed. The new definition is *Group quarters unit, not specified above*.
Code 11 has been eliminated.

The following changes apply to POPSIZE:

Code 4 (*75-329.9 thousand*) has been changed. The new definition is *125-329.9 thousand*.
Code 5 (*Less than 75 thousand*) has been changed. The new definition is *less than 125 thousand*.

- 5) The following variables in the FMLY files have attribute changes.

EDUC_REF (*CHAR(1)*) has been changed. The new attribute is *CHAR(2)*.
EDUCA2 (*CHAR(1)*) has been changed. The new attribute is *CHAR(2)*.
FS_DATE 1 through FS_DATE8 (*NUM(6)*) has been changed. The new attribute is *NUM(8)*.
STRTYEAR (*CHAR(2)*) has been changed. The new attribute is *CHAR(4)*.

- 6) The following variables have been deleted from the MEMB files.

COMPLET Was highest school grade completed?
COMPLET_ COMPLET flag

7) The variable EDUCA in the MEMB files have the following code and code definition changes:

The codes eliminated are:

- 00 Never attended school
- 01-12 First grade through twelfth grade or equivalent
- 21 First year of college or equivalent
- 22 Second year of college or equivalent
- 23 Third year of college or equivalent
- 24 Fourth year of college or equivalent
- 25 One year of graduate school
- 26 Two or more years of graduate school

The new codes that apply are:

- 00 Never attended school
- 01-11 First through eleventh grade
- 38 Twelfth grade - no degree
- 39 High school graduate
- 40 Some college - no degree
- 41 AA degree (occupational/vocational)
- 42 AA degree (academic)
- 43 Bachelors degree
- 44 Masters degree
- 45 Professional degree
- 46 Doctorate degree

8) The following UCC has been added to the EXPN files.

310334 Satellite dishes

9) The following variable in the EXPN files have attribute changes.

QREDATE (*CHAR(8)*) has been changed. The new attribute is *CHAR(10)*.

10) The following UCC's have undergone content changes.

200310 Wine at Home
- Nonalcoholic wine is now mapped to 200310.

180220 Frozen /Prepared Food Other than Meals
- Frozen buffalo wings is now mapped to 180220.

180710 Miscellaneous Prepared Food
- Bottled/Canned Buffalo Wings is now mapped to 180710.

340120 Delivery Services
- Fax services is now mapped to 340120.

620911 Miscellaneous Fees, Parimutuel Losses
- Lottery tickets is now mapped to 620911.

11) The following PUBFLAG value changes begin in Q19961

<u>UCC</u>	<u>New PUBFLAG values</u>
190902	2
280110	1
280120	1
280130	1
280230	1
310110	1
320150	2
320210	2
320220	1
320320	2
320420	2
320902	1
320903	1
340520	2
360350	2
360901	2
370110	1
370311	1
370312	1
370313	1
380430	2
390120	2
410120	2
410901	2
420110	2
420120	2
430110	2
430120	2
440120	2
440210	2
480213	1
550340	1
600210	2
600410	2
600420	2
610110	2
610120	2
650110	1
690114	1

III. FILE INFORMATION

Commencing with the 1996 Diary release, the public use microdata will consist of ASCII files and SAS data sets on CD-ROM; data will no longer be released on tapes.

The 1996 Diary release contains four sets of Diary data files (FMLY, MEMB, EXPN, DTAB) and four processing files. The FMLY, MEMB, EXPN, and DTAB files are organized by the quarter of the calendar year in which the data were collected. (See Section V.A.1.b for description of calendar and

collection years.) There are four quarterly data sets for each of the following files: a consumer unit (CU) characteristics, income, and summary level expenditure file (FMLY), a member characteristics and income file (MEMB), a detailed expenditure file (EXPN), and an income file (DTAB).

The four processing files are used to enhance computer processing and tabulation of data, and to provide descriptive information on item codes. Processing files are as follows: a sample table aggregation file (AGG), a sample table label file (LAB), a Universal Classification Codes file (UCC), and a file (SAMPL) containing the sample program (Section VII.A.) The processing files are further explained in Section III.E.5

A file containing this complete documentation is included on the X:\Document directory of the CD-ROM as an Adobe Acrobat PDF file and is named *Drydoc96.pdf*. The appropriate Adobe Acrobat Reader is required to read and print this file. The reader is provided in the X:\Acroread subdirectory of the compact disk and can be loaded onto your system by following the guidelines in the *Readme.1st* file on the root directory. Adobe Reader is a shareware product.

Note that the variable NEWID, the CU's identification number, is the common variable among files by which matching is done.

Logical record lengths of data and processing files are as follows:

FMLY	LRECL =1549
MEMB	LRECL = 247
EXPN	LRECL = 40
DTAB	LRECL = 28
AGG	LRECL = 80
LAB	LRECL = 80
UCC	LRECL = 80
DOC	LRECL = 80

A. DATA SET NAMES

The ASCII data set names are as follows:

```
X:\DIARY96\FMLYD961.txt (Diary FMLY file for first quarter, 1996)
X:\DIARY96\MEMBD961.txt (Diary MEMB file for first quarter, 1996)
X:\DIARY96\EXPND961.txt (Diary EXPN file for first quarter, 1996)
X:\DIARY96\DTABD961.txt (Diary DTAB file for first quarter, 1996)
X:\DIARY96\FMLYD962.txt (etc.)
X:\DIARY96\MEMBD962.txt
X:\DIARY96\EXPND962.txt
X:\DIARY96\DTABD962.txt
X:\DIARY96\FMLYD963.txt
X:\DIARY96\MEMBD963.txt
X:\DIARY96\EXPND963.txt
X:\DIARY96\DTABD963.txt
X:\DIARY96\FMLYD964.txt
X:\DIARY96\MEMBD964.txt
X:\DIARY96\EXPND964.txt
X:\DIARY96\DTABD964.txt
X:\DIARY96\AGGD96.txt
X:\DIARY96\LABELD96.txt
X:\DIARY96\UCCD96.txt
X:\DIARY96\DOCD96.txt
```


where "X" references the designated drive for your CD.

The SAS data set names are as follows:

```
X:\DIARY96\FMLD961.sd2  (Diary FMLY file for first quarter, 1996)
X:\DIARY96\MEMD961.sd2  (Diary MEMB file for first quarter, 1996)
X:\DIARY96\EXPD961.sd2  (Diary EXPN file for first quarter, 1996)
X:\DIARY96\DTBD961.sd2  (Diary DTAB file for first quarter, 1996)
X:\DIARY96\FMLD962.sd2  (etc.)
X:\DIARY96\MEMD962.sd2
X:\DIARY96\EXPD962.sd2
X:\DIARY96\DTBD962.sd2
X:\DIARY96\FMLD963.sd2
X:\DIARY96\MEMD963.sd2
X:\DIARY96\EXPD963.sd2
X:\DIARY96\DTBD963.sd2
X:\DIARY96\FMLD964.sd2
X:\DIARY96\MEMD964.sd2
X:\DIARY96\EXPD964.sd2
X:\DIARY96\DTBD964.sd2
X:\DIARY96\AGGD96.sd2
X:\DIARY96\LABELD96.sd2
X:\DIARY96\UCCD96.sd2
X:\DIARY96\DOCD96.sd2
```

B. RECORD COUNTS PER QUARTER

The number of records in each data set are as follows:

<u>ASCII data set</u>	<u>SAS data set</u>	<u>Record Count</u>
FMLYD961.txt	FMLD961.sd2	2,135
MEMBD961.txt	MEMD961.sd2	5,430
EXPND961.txt	EXPD961.sd2	89,058
DTABD961.txt	DTBD961.sd2	33,716
FMLYD962.txt	FMLD962.sd2	2,481
MEMBD962.txt	MEMD962.sd2	6,436
EXPND962.txt	EXPD962.sd2	107,656
DTABD962.txt	DTBD962.sd2	39,656
FMLYD963.txt	FMLD963.sd2	2,592
MEMBD963.txt	MEMD963.sd2	6,691
EXPND963.txt	EXPD963.sd2	111,359
DTABD963.txt	DTBD963.sd2	41,508
FMLYD964.txt	FMLD964.sd2	3,568
MEMBD964.txt	MEMD964.sd2	9,155
EXPND964.txt	EXPD964.sd2	151,625
DTABD964.txt	DTBD964.sd2	55,928

C. DATA FLAGS:

Data fields on the FMLY and MEMB files are explained by flag variables following the data field. The flag variables names are derived from the names of the data fields they reference. In general the rule is to add an underscore to the last position of the data field name (for example WAGEX becomes

WAGEX_). However, if the data field name is eight characters in length, then the fifth position is replaced with an underscore. If this fifth position is already an underscore, then the fifth position is changed to a zero (for example EDUC_REF becomes EDUC0REF).

The flag values are defined as follows:

A flag value of "A" indicates a valid blank; that is, a blank field where a response is not anticipated.

A flag value of "B" indicates a blank resulting from an invalid nonresponse; that is, a nonresponse that is not consistent with other data reported by the CU.

A flag value of "C" refers to a blank resulting from a "don't know", refusal, or other type of nonresponse.

A flag value of "D" indicates that the characteristics or weight factor field contains a valid or good data value.

A flag value of "T" indicates topcoding has been applied to the data field.

A flag value of "R" for recode has been created for the variable STATE_ in 1996. Commencing with the 1996 sample design, some Primary Sampling Units in some states are given "false" STATE codes for nondisclosure reasons. CUs with STATE_='R' (for recode) indicate that not all CUs with that particular STATE code are from that state. See section on topcoding for more detail.

D. FILE NOTATION

Every record from each data file includes the variable NEWID, the CU's unique identification number, which can be used to link records of one CU from several files, for example FMLY and MEMB, across all quarters in which they participate.

Data fields for variables on the microdata files have either numeric or character values. The format column in each data file distinguishes whether a variable is numeric (NUM) or character (CHAR) and shows the number of field positions the variable occupies. Variables which include decimal points are formatted as NUM(t,r) where t is the total number of positions occupied, and r is the number of places to the right of the decimal.

Besides format, this documentation's detailed variable listings give an item description, questionnaire source, identification of codes where applicable, and start position for each variable. The source, which identifies where the data for that variable is collected on the characteristics questionnaire, is listed beneath the variable description and has a format such as "S04B 2b", which denotes Section 4, Part B, Question 2b of the characteristics questionnaire.

A star (*) is shown in front of new variables, those which have changed in format or definition, and those which have been deleted. New variables are added to the end of the files.

Some variables require special notation. The following notation is used throughout the documentation for all files:

*D(Yxxq) identifies a variable which is deleted as of the quarterly file indicated. The year and quarter are identified by the 'xx' and 'q' respectively. For example, the notation *D(Y961) indicates the variable is deleted starting with the data file of the first quarter of 1996.

*N(Yxxq) identifies a variable which is added as of the quarterly file indicated. The year and quarter are identified by the 'xx' and 'q' for new variables in the same way as for deleted variables.

*L indicates that the variable can contain negative values.

E. DETAILED VARIABLE DESCRIPTIONS

1. CONSUMER UNIT (CU) CHARACTERISTICS AND INCOME FILE (FMLY)

The "FMLY" file, also referred to as the "Consumer Unit Characteristics and Income" file, contains CU characteristics, CU income, characteristics and earnings of the reference person and of the spouse. The file includes weights needed to calculate population estimates and variances. (See Sections V. and VI.)

Summary expenditure variables in this file can be used to derive estimates for broad consumption categories. These variables aggregate expenditures to match the level of detail published in previous Diary News Releases.

When there is a valid nonresponse, or where nonresponse occurs and there is no imputation, there will be missing values. The type of nonresponse is explained by associated data flag variables described in Section III.C. DATA FLAGS.

a. CU AND DIARY IDENTIFIERS

VARIABLE	ITEM DESCRIPTION	START POSITION	FORMAT
NEWID	CU identification number. Digits 1-7 (CU sequence number, 0000001 through 9999999) uniquely identifies the CU. Digit 8 is the week number, 1 or 2	1	NUM(8)
	BLS derived		
HH_CU_Q	Count of CUs in this household	1507	NUM(2)
	BLS derived		
HH_CU_Q_		1509	CHAR(1)
HHID	Identifier for household with more than one CU. Household with only one CU will be set to missing.	1510	NUM(3)
	BLS derived		
HHID_		1513	CHAR(1)

WEEKI	Week of the Diary CODED 1 First week Diary 2 Second week Diary Census derived	656	CHAR(1)
WEEKI_		657	CHAR(1)
WEEKN	Number of Diary weeks surveyed, 1 or 2 BLS derived	658	NUM(1)
STRTDAY	Start day of this Diary week Cover 19	625	CHAR(2)
STRTMNTH	Start month of this Diary week Cover 19	627	CHAR(2)
STRTYEAR	Start year of this Diary week Cover 19	629	CHAR(4)
PICK_UP	Interview status at pick-up CODED Interview status at pick-up 01 Diary placed or completed 03 Temporarily absent during entire reference period Cover 20	559	CHAR(2)

b. CU CHARACTERISTICS

VARIABLE	ITEM DESCRIPTION	START POSITION	FORMAT
*REGION	Region CODED 1 Northeast 2 Midwest 3 South 4 West BLS derived	580	CHAR(1)
REGION_		581	CHAR(1)
BLS_URBN	Urban/Rural CODED 1 Urban 2 Rural BLS derived	42	CHAR(1)

*POPSIZE Population size of the PSU
 CODED 564 CHAR(1)
 1 More than 4 million
 2 1.20-4 million
 3 0.33-1.19 million
 4 125 - 329.9 thousand
 5 Less than 125 thousand

BLS derived

SMSASTAT Does CU reside inside an MSA?
 CODED 606 CHAR(1)
 1 Yes, resides inside an MSA
 2 No, resides outside an MSA

BLS derived

* STATE State identifier (see Section IV.A. and Section X.D. for important
 information) 1518 CHAR(2)

01	Alabama	*28	Mississippi
02	Alaska	**29	Missouri
^{RR} 04	Arizona	31	Nebraska
*05	Arkansas	^R 32	Nevada
**06	California	^R 33	New Hampshire
08	Colorado	34	New Jersey
09	Connecticut	*35	New Mexico
10	Delaware	^{RR} **36	New York
^R 11	District of Columbia	**37	North Carolina
**12	Florida	^{RR} 39	Ohio
**13	Georgia	**40	Oklahoma
15	Hawaii	**41	Oregon
16	Idaho	42	Pennsylvania
**17	Illinois	45	South Carolina
^{RR} **18	Indiana	*46	South Dakota
*19	Iowa	**47	Tennessee
**20	Kansas	48	Texas
21	Kentucky	49	Utah
22	Louisiana	50	Vermont
^R *23	Maine	**51	Virginia
24	Maryland	**53	Washington
25	Massachusetts	^R 54	West Virginia
**26	Michigan	55	Wisconsin
**27	Minnesota		

* indicates that the STATE code has been suppressed for all sampled CUs in that state (STATE_ = 'T' for all observations).

** indicates that the STATE code has been suppressed for some sampled CUs in that state (STATE_ = 'T' for some observations).

^R indicates that either all observations from this state have been re-coded or all strata¹ of observations from this state include "re-codes" from other states.

^{RR} indicates that either some observations from this state have been re-coded or at least one stratum¹ of observations from this state includes "re-codes" from other states.

^{R*} indicates that the STATE code has been suppressed and, either STATE has been re-coded or that state includes "re-codes" from other states for all strata¹ from that state.

RR** indicates that the STATE code has been suppressed and, either STATE has been re-coded or that state includes “re-codes” from other states for some strata¹ from that state.

¹ A STATE stratum is a unique POPSIZE and BLS_URBN combination.

States not listed are not in the CE sample.

* STATE_	*N(Y961)	1520	CHAR(1)
CUTENURE	Housing tenure CODED 1 Owned with mortgage 2 Owned without mortgage 3 Owned mortgage not reported 4 Rented 5 Occupied without payment of cash rent 6 Student housing BLS derived	43	CHAR(1)
CUTE_URE		44	CHAR(1)
FAM_SIZE	Number of members in CU BLS derived	78	NUM(2)
FAM__IZE		80	CHAR(1)
PERSLT18	Number of children less than 18 in CU BLS derived	544	NUM(2)
PERS_T18		546	CHAR(1)
PERSOT64	Number of persons over 64 in CU BLS derived	547	NUM(2)
PERS_T64		549	CHAR(1)
CHILDAGE	Age of children of reference person CODED 0 No children 1 All children less than 6 2 Oldest child between 6 and 11 and at least one child less than 6 3 All children age 6 to 11 4 Oldest child between 12 and 17 and at least one child less than 12 5 All children between 12 and 17 6 Oldest child greater than 17 and at least one child less than 17 7 All children greater than 17 BLS derived	1514	CHAR(1)

CHIL__AGE		1515	CHAR(1)
FAM__TYPE	CU type is based on relationship of members to reference person. "Own" children include blood-related sons and daughters, step children and adopted children. CODED 1 Husband and wife (H/W) only 2 H/W, own children only, oldest child < 6 3 H/W, own children only, oldest child > 5, <= 17 4 H/W, own children only, oldest child > 17 5 All other H/W CUs 6 One parent, male, own children only, at least one child age < 18 7 One parent, female, own children only, at least one child age < 18 8 Single persons 9 Other CUs BLS derived	81	CHAR(1)
FAM__YPE		82	CHAR(1)
NO__EARNR	Number of earners BLS derived	471	NUM(2)
NO__E__RNR		473	CHAR(1)
EARNCOMP	Composition of earners CODED 1 Reference person only 2 Reference person and spouse 3 Reference person, spouse, and others 4 Reference person and others 5 Spouse only 6 Spouse and others 7 Others only 8 No earners BLS derived	57	CHAR(1)
EARN__OMP		58	CHAR(1)
VEHQ	Number of vehicles S02 4B	653	NUM(2)
VEHQ__		655	CHAR(1)
INCLASS	Income class of CU based on income before taxes(Codes 01 through 09 are for CUs considered complete reporters of income) CODED 01 Less than \$5,000 02 \$5,000 to \$9,999	1516	CHAR(2)

03 \$10,000 to \$14,999
 04 \$15,000 to \$19,999
 05 \$20,000 to \$29,999
 06 \$30,000 to \$39,999
 07 \$40,000 to \$49,999
 08 \$50,000 to \$69,999
 09 \$70,000 and over
 10 Incomplete income reported

BLS derived

RESPSTAT	Completeness of income response CODED	582	CHAR(1)
	1 Complete income respondent		
	2 Incomplete income respondent		

BLS derived

RESP_TAT		583	CHAR(1)
----------	--	-----	---------

INC_RNKU	Weighted cumulative percent income ranking of CU to total population. Ranking based on total income before taxes. Rank of incomplete income reporters is set to zero.	395	NUM(9,7)
----------	---	-----	----------

BLS derived

INC__NKU		404	CHAR(1)
----------	--	-----	---------

* POVERTY	Is CU income below current year's poverty threshold? CODED	1548	CHAR(1)
	1 Yes, below the current year's poverty threshold		
	2 No, not below current year's poverty threshold		
	*N(961)		

BLS derived

*POVERTY_	*N(961)	1549	CHAR(1)
-----------	----------------	------	---------

c. CHARACTERISTICS OF REFERENCE PERSON AND SPOUSE

VARIABLE	ITEM DESCRIPTION	START POSITION	FORMAT
AGE_REF	Age of reference person	36	NUM(2)
	S01 9		
AGE_REF_		38	CHAR(1)
REF_RACE	Race of reference person CODED	578	CHAR(1)
	1 White		

	2 Black		
	3 American Indian, Aleut, Eskimo		
	4 Asian or Pacific Islander		
	5 Other		
	S01 10		
REF__ACE		579	CHAR(1)
SEX_REF	Sex of reference person	602	CHAR(1)
	CODED		
	1 Male		
	2 Female		
	S01 6		
SEX_REF_		603	CHAR(1)
MARITAL1	Marital status of reference person	469	CHAR(1)
	CODED		
	1 Married		
	2 Widowed		
	3 Divorced		
	4 Separated		
	5 Never married		
	S01 12		
MARI_AL1		470	CHAR(1)
ORIGIN1	Origin or ancestry of reference person	495	CHAR(1)
	CODED		
	1 European:		
	German		
	Italian		
	Irish		
	French		
	Polish		
	Russian		
	English		
	Scottish		
	Dutch		
	Swedish		
	Hungarian		
	2 Spanish:		
	Mexican American		
	Chicano Mexican		
	Puerto Rican		
	Cuban		
	Central or South American		
	Other Spanish		
	3 Afro-American		
	4 Other and Don't know		
	S01 11		

ORIGIN1_		496	CHAR(1)
* EDUC_REF	Education of reference person CODED 00 Never attended school 10 First through eighth grade 11 Ninth through twelve grade (no H.S. diploma) 12 High school graduate 13 Some college, less than college graduate 14 AA degree (occupational/vocational or academic) 15 Bachelors degree 16 Masters degree 17 Professional/doctorate degree BLS derived	68	CHAR(2)
EDUC0REF		70	CHAR(1)
AGE2	Age of spouse S01 9	39	NUM(2)
AGE2_		41	CHAR(1)
RACE2	Race of spouse CODED - same as REF_RACE S01 10	574	CHAR(1)
RACE2_		575	CHAR(1)
SEX2	Sex of spouse CODED - same as SEX_REF S01 6	604	CHAR(1)
SEX2_		605	CHAR(1)
ORIGIN2	Origin or ancestry of spouse CODED - same as ORIGIN1 S01 12	497	CHAR(1)
ORIGIN2_		498	CHAR(1)
*EDUCA2	Education of spouse CODED - same as EDUC_REF BLS derived	71	CHAR(2)
EDUCA2_		73	CHAR(1)

d. **WORK EXPERIENCE OF REFERENCE PERSON AND SPOUSE**

VARIABLE	ITEM DESCRIPTION	START POSITION	FORMAT
WK_WRKD1	Number of weeks worked (including paid vacation and paid sick leave) by reference person full or part time in last 12 months S04A 2	672	NUM(2)
WK_W_KD1		674	CHAR(1)
HRSPRWK1	Number of hours usually worked per week by reference person S04A 3	387	NUM(3)
HRSP_WK1		390	CHAR(1)
OCCULIS1	The job in which reference person received the most earnings during the past 12 months best in fits the following category CODED Manager, professional 01 Administrator, manager 02 Teacher 03 Professional Administrative support, technical, sales 04 Administrative support, including clerical 05 Sales, retail 06 Sales, business goods and services 07 Technician Service 08 Protective service 09 Private household service 10 Other service Operator, assembler, laborer 11 Machine operator, assembler, inspector 12 Transportation operator 13 Handler, helper, laborer Precision production, craft, repair 14 Mechanic, repairer, precision production 15 Construction, mining Farming, forestry, fishing 16 Farming 17 Forestry, fishing, groundskeeping Armed forces 18 Armed forces S04A 4a	561	CHAR(2)
OCCU_IS1		563	CHAR(1)
EMPLTYP1	Employer from which reference person received the most earnings in past 12 months CODED 1 Private company, business, or individual 2 Federal government	74	CHAR(1)

- 3 State government
- 4 Local government
- 5 Self-employed in own business, professional practice, or farm
- 6 Family business or farm, working without pay

S04A 4b

EMPL_YP1		75	CHAR(1)
WHYNWRK1	Reason reference person did not work last 12 months CODED	668	CHAR(1)
	1 Retired		
	2 Taking care of home/family		
	3 Going to school		
	4 Ill, disabled, unable to work		
	5 Unable to find work		
	6 Doing something else		

S04A 5

WHYN_RK1		669	CHAR(1)
WK_WRKD2	Number of weeks worked (including paid vacation and paid sick leave) by spouse full or part time in last 12 months	675	NUM(2)

S04A 2

WK_W_KD2		677	CHAR(1)
HRSPRWK2	Number of hours usually worked per week by spouse	391	NUM(3)

S04A 3

HRSP_WK2		394	CHAR(1)
OCCULIS2	Job in which spouse received the most earnings during the past 12 months CODED - same as OCCULIS1	492	CHAR(2)

S04A 4a

OCCU_IS2		494	CHAR(1)
EMPLTYP2	Employer from which spouse received the most earnings the past 12 months CODED - Same as EMPLTYP1	76	CHAR(1)

S04A 4b

EMPL_YP2		77	CHAR(1)
WHYNWRK2	Reason spouse did not work last 12 months CODED - same as WHYNWRK1	670	CHAR(1)

S04A 5

WHYN_RK2		671	CHAR(1)
OCCEXPX	Annual occupational expenses such as union dues, tools, uniforms, business or professional association dues, licenses or permits in the last 12 months	483	NUM(8)
	S04B 5		
OCCE_PNX		491	CHAR(1)

e. **INCOME**

VARIABLE	ITEM DESCRIPTION	START POSITION	FORMAT
FINCBEFX	Amount of CU income before taxes in past 12 months (UNEMPX + WRKRSX + WELFRX + INTX +DIVX + PENSIONX + ROOMX + OTHRNTX + CHDOTX + ALIOTHX + OTHINX + JFS_AMT + FWAGEX + FBSNSX + FFARMX + FSS_RRX + FSUPPX) *L BLS derived	139	NUM(8)
FINC_EFX		147	CHAR(1)
FINCAFTX	Amount of CU income after taxes in past 12 months (FINCBEFX - PERSTAX) *L BLS derived	130	NUM(8)
FINC_FTX		138	CHAR(1)
EARNX	Amount of earned income before taxes by CU in past 12 months (FWAGEX + FBSNSX + FFARMX) *L BLS derived	59	NUM(8)
EARNX_		67	CHAR(1)
NONERNX	Amount of CU income other than earnings before taxes in past 12 months (FSS_RRX + FSUPPX + UNEMPX + WRKRSX + WELFRX + INTX + DIVX + PENSIONX + ROOMX + OTHRNTX + CHDOTX + ALIOTHX + OTHINX + JFS_AMT) *L BLS derived	474	NUM(8)
NONERNX_		482	CHAR(1)
FWAGEX	Amount of wage and salary income before deductions received	378	NUM(8)

	by all CU members in past 12 months (Sum WAGEX from MEMB file for all CU members)		
	BLS derived		
FWAGEX_		386	CHAR(1)
FBSNSX	Amount of income or loss from non-farm business, partnership or professional practice received by all CU members in past 12 months (Sum BSNSX from MEMB file for all CU members) *L	83	NUM(8)
	BLS derived		
FBSNSX_		91	CHAR(1)
FFARMX	Amount of income or loss from own farm received by all CU members in past 12 months *L (Sum FARMX from MEMB file for all CU members) BLS derived	103	NUM(8)
FFARMX_		111	CHAR(1)
FSS_RRX	Amount of income from Social Security and Railroad Retirement prior to deductions for medical insurance and Medicare received by all CU members in past 12 months (Sum SOCRRX from MEMB file for all CU members)	351	NUM(8)
	BLS derived		
FSS_RRX_		359	CHAR(1)
FSUPPX	Amount of income from all sources of Supplemental Security Income received by all CU members in past 12 months (Sum SUPPX from MEMB file for all CU members)	369	NUM(8)
	BLS derived		
FSUPPX_		377	CHAR(1)
UNEMPX	Amount of income from unemployment compensation received by CU in past 12 months	644	NUM(8)
	S04B 1a		
UNEMPX_		652	CHAR(1)
WRKRSX	Amount of income from worker's compensation and veterans' payments, including education benefits but excluding military retirement, received by CU in past 12 months	678	NUM(8)
	S04B 1b		
WRKRSX_		686	CHAR(1)

WELFRX	Amount of income from public assistance or welfare including job training grants such as Job Corps received by CU in past 12 months	659	NUM(8)
	S04B 1c		
WELFRX_		667	CHAR(1)
INTX	Amount of income from interest on savings accounts or bonds received by CU in past 12 months	414	NUM(8)
	S04B 1d		
INTX_		422	CHAR(1)
DIVX	Amount of regular income from dividends royalties, estates, or trusts received by CU in past 12 months	48	NUM(8)
	S04B 1e		
DIVX_		56	CHAR(1)
PENSIONX	Amount of income from pensions or annuities from private companies, military or government, IRA, or Keogh received by CU in past 12 months	535	NUM(8)
	S04B 1f		
PENS_ONX		543	CHAR(1)
ROOMX	Amount of net income or loss from roomers or boarders received by CU in past 12 months	584	NUM(8)
	*L		
	S04B 1g(1)		
ROOMX_		592	CHAR(1)
OTHRNTX	Amount of net income or loss from other rental units received by CU in past 12 months	526	NUM(8)
	*L		
	S04B 1g(2)		
OTHRNTX_		534	CHAR(1)
OTHINX	Amount of other money income including money from care of foster children, cash scholarships and fellowships, or stipends not based on working, received by CU in past 12 months	499	NUM(8)
	S04B 2c		
OTHINX_		507	CHAR(1)
CHDOTHX	Total amount received by all CU members in last 12 months for	1521	NUM(8)

other child support

S04B 1h(2)

CHDOTHX_		1529	CHAR(1)
----------	--	------	---------

ALIOTHX	Total amount received from regular contributions by all CU members	1530	NUM(8)
---------	--	------	--------

S04B 1i(2)

ALIOTHX_		1538	CHAR(1)
----------	--	------	---------

JFS_AMT	Annual value of food stamps received by CU JFS_AMT = 12 X sum of (FS_AMT1 ... FS_AMT8) NOTE: JFS_AMT is a component of FINCBEFX, NONERNX, and FINCAFTX	423	NUM(8)
---------	---	-----	--------

BLS derived

JFS_AMT_		431	CHAR(1)
----------	--	-----	---------

f. OTHER MONEY RECEIPTS

VARIABLE	ITEM DESCRIPTION	START POSITION	FORMAT
OTHRECX	Amount of other money receipts received by CU in past 12 months (LUMPX + SALEX + SSREFX + INSREFX + PTAXREF)	508	NUM(8)
	BLS derived		
OTHRECX_		516	CHAR(1)
LUMPX	Amount of lump sum receipts from estates, trusts, royalties, alimony, prizes, games of chance, or persons outside CU received by CU in past 12 months	460	NUM(8)
	S04B 2a		
LUMPX_		468	CHAR(1)
CHDLMPX	Total amount received by all CU members for a lump sum child support payment in last 12 months	1539	NUM(8)
	S04B 1h(1)		
CHDLMPX_		1547	CHAR(1)
SALEX	Amount from sale of household furnishings, equipment, clothing, jewelry, pets or other belongings, excluding sale of vehicles or property received by CU in past 12 months	593	NUM(8)
	S04B 2b		

SALEX_		601	CHAR(1)
SSREFX	Amount of refunds from overpayment on Social Security received by CU in past 12 months	607	NUM(8)
	S04B 3c		
SSREFX_		615	CHAR(1)
INSREFX	Amount of refunds from insurance policies received by CU in past 12 months	405	NUM(8)
	S04B 3d		
INSREFX_		413	CHAR(1)
PTAXREFX	Amount of refunds from property taxes received by CU in past 12 months	565	NUM(8)
	S04B 3e		
PTAX_EFX		573	CHAR(1)

g. TAXES

VARIABLE	ITEM DESCRIPTION	START POSITION	FORMAT
PERSTAX	Amount of personal taxes paid by CU in past 12 months (ADDFEDX + ADDSTAX + ADDOTHX + FFEDTXX + FSTATXX + TAXPROPX) - (FEDREFX + STATREFX + OTHREFX) *L BLS derived	550	NUM(8)
PERSTAX_		558	CHAR(1)
TAXPROPX	Amount of personal property taxes paid but not reported elsewhere by CU in past 12 months S04B 4c	633	NUM(8)
TAXP_OPX		641	CHAR(1)
FFEDTXX	Amount of Federal income tax deducted from last pay for all CU members ((FEDTXX/GROSPAYX) x WAGEX) BLS derived	112	NUM(8)
FFEDTXX_		120	CHAR(1)
ADDFEDX	Amount of Federal income tax paid in addition to that withheld from earnings by all CU members	9	NUM(8)

S04B 4a			
ADDFEDX_		17	CHAR(1)
FEDREFX	Amount of refund from Federal income tax received by CU in past 12 months	94	NUM(8)
S04B 3a			
FEDREFX_		102	CHAR(1)
FSTATXX	Amount of state and local income taxes deducted for all CU members ((STATXX/GROSPAYX) x WAGEX)	360	NUM(8)
BLS derived			
FSTATXX_		368	CHAR(1)
ADDSTAX	Amount of state and local income taxes paid in addition to that withheld from earnings by all CU members in past 12 months	27	NUM(8)
S04B 4b			
ADDSTAX_		35	CHAR(1)
STATREFX	Amount of refund from state and local income taxes received by CU in past 12 months	616	NUM(8)
S04B 3b			
STAT_EFX		624	CHAR(1)
ADDOTHX	Amount of other taxes paid but not reported elsewhere by CU in past 12 months	18	NUM(8)
S04B 4d			
ADDOTHX_		26	CHAR(1)
OTHREFX	Amount of refunds from other sources, including any other taxes, received by CU in past 12 months	517	NUM(8)
S04B 3f			
OTHREFX_		525	CHAR(1)

h. RETIREMENT AND PENSION DEDUCTIONS

VARIABLE	ITEM DESCRIPTION	START POSITION	FORMAT
FJSSDEDX	Estimated amount of income contributed to Social Security by all CU members in past 12 months (Sum JSSDEDX from MEMB file for all CU members)	168	NUM(8)
	BLS derived		
FJSS_EDX		176	CHAR(1)
FRRX	Amount of Railroad Retirement deducted from last pay for all CU members (Sum ((RRX/GROSPAYX) x WAGEX) from MEMB file for all CU members)	195	NUM(8)
	BLS derived		
FRRX_		203	CHAR(1)
FGVX	Amount of government retirement deducted from last pay for all CU members (Sum ((GVX/GROSPAYX) x WAGEX) from MEMB file for all CU members)	121	NUM(8)
	BLS derived		
FGVX_		129	CHAR(1)
FPVTX	Amount of private pension fund deducted from last pay for all CU members (sum ((PVTX/GROSPAYX) x WAGEX) from MEMB file, for all CU members)	177	NUM(8)
	BLS derived		
FPVTX_		185	CHAR(1)
FIRAX	Amount deducted for individual retirement plan, such as an IRA or Keogh, for all CU members (sum IRAX from MEMB file for all CU members)	159	NUM(8)
	BLS derived		
FIRAX_		167	CHAR(1)

i. **FOOD STAMPS**

NOTE: JFS_AMT is in e. INCOME above.

VARIABLE	ITEM DESCRIPTION	START POSITION	FORMAT
REC_FS	Did any members of CU receive Food Stamps, during past 12 months? CODED 1 Yes, member in CU received Food Stamps 2 No, no member in CU received Food Stamps S04B 8a	576	CHAR(1)
REC_FS_		577	CHAR(1)
FD_STMPS	Did any members of CU receive Food Stamps, during past month? CODED 1 Yes 2 No S04B 9a	92	CHAR(1)
FD_S_MPS		93	CHAR(1)
FS_MTHI	Number of months Food Stamps received during past 12 months S04B 8b	348	NUM(2)
FS_MTHI_		350	CHAR(1)
FS_AMT1	Dollar value of Food Stamps received for the first entry during the past month S04B 9c	204	NUM(8)
FS_AMT1_		212	CHAR(1)
FS_AMT2	Dollar value of Food Stamps received for the second entry during the past month S04B 9c	213	NUM(8)
FS_AMT2_		221	CHAR(1)
FS_AMT3	Dollar value of Food Stamps received for the third entry during the past month S04B 9c	222	NUM(8)
FS_AMT3_		230	CHAR(1)
FS_AMT4	Dollar value of Food Stamps received for the fourth entry during	231	NUM(8)

	the past month		
	S04B 9c		
FS_AMT4_		239	CHAR(1)
FS_AMT5	Dollar value of Food Stamps received for the fifth entry during the past month	240	NUM(8)
	S04B 9c		
FS_AMT5_		248	CHAR(1)
FS_AMT6	Dollar value of Food Stamps received for the sixth entry during the past month	249	NUM(8)
	S04B 9c		
FS_AMT6_		257	CHAR(1)
FS_AMT7	Dollar value of Food Stamps received for the seventh entry during the past month	258	NUM(8)
	S04B 9c		
FS_AMT7_		266	CHAR(1)
FS_AMT8	Dollar value of Food Stamps received for the eighth entry during the past month	267	NUM(8)
	S04B 9c		
FS_AMT8_		275	CHAR(1)
* FS_DATE1	Month, day, and year Food Stamps were received for the first entry during the past month	276	NUM(8)
	S04B 9b		
FS_D_TE1		284	CHAR(1)
* FS_DATE2	Month, day, and year Food Stamps were received for the second entry during the past month	285	NUM(8)
	S04B 9b		
FS_D_TE2		293	CHAR(1)
* FS_DATE3	Month, day, and year Food Stamps were received for the third entry during the past month	294	NUM(8)
	S04B 9b		
FS_D_TE3		302	CHAR(1)
* FS_DATE4	Month, day, and year Food Stamps were received for the fourth	303	NUM(8)

entry during the past month

S04B 9b

FS_D_TE4		311	CHAR(1)
* FS_DATE5	Month, day, and year Food Stamps were received for the fifth entry during the past month	312	NUM(8)
	S04B 9b		
FS_D_TE5		320	CHAR(1)
* FS_DATE6	Month, day, and year Food Stamps were received for the sixth entry during the past month	321	NUM(8)
	S04B 9b		
FS_D_TE6		329	CHAR(1)
* FS_DATE7	Month, day, and year Food Stamps were received for the seventh entry during the past month	330	NUM(8)
	S04B 9b		
FS_D_TE7		338	CHAR(1)
* FS_DATE8	Month, day, and year Food Stamps were received for the eighth entry during the past month	339	NUM(8)
	S04B 9b		
FS_D_TE8		347	CHAR(1)

j. **FREE MEALS AND GROCERIES**

VARIABLE	ITEM DESCRIPTION	START POSITION	FORMAT
FREEMLX	Weekly dollar value of free meals received as part of pay by any CU members during the past 12 months	186	NUM(8)
	S04B 6b		
FREEMLX_		194	CHAR(1)
JGROCYMV	Monthly expenditure for grocery store expenditures	446	NUM(6)
	S02 3a Computation		
JGRO_YMV		452	CHAR(1)
JGROCYWK	Weekly expenditure for grocery store purchases	453	NUM(6)
	S02 3a Computation		

JGRO_YWK		459	CHAR(1)
JGRCFDMV	Monthly grocery expenditures for food and non- alcoholic beverages purchased at grocery stores	432	NUM(6)
	S02 3a-3b Computation		
JGRC_DMV		438	CHAR(1)
JGRCFDWK	Weekly grocery purchases for food and non- alcoholic beverages purchased at grocery stores	439	NUM(6)
	S02 3a-3b Computation		
JGRC_DWK		445	CHAR(1)

k. **HOUSING STRUCTURE**

VARIABLE	ITEM DESCRIPTION	START POSITION	FORMAT
*DESCRIP	Housing unit description CODED 01 House, apartment, flat 02 Housing unit in nontransient hotel, motel, etc. 03 Housing unit, permanent in transient hotel, motel, etc. 04 Housing unit, in rooming house 05 Mobile home or trailer with NO permanent room added 06 Mobile home or trailer with one or more permanent rooms added 07 Housing unit not specified above 08 Quarters not housing unit in rooming or boarding house 09 Student quarters in college dormitory 10 Group quarters unit, not specified above	45	CHAR(2)
DESCRIP_	Cover 13c and 13d	47	CHAR(1)
TYPOWND	Are living quarters owned by regular ownership or as a condominium or cooperative? CODED 1 Regular ownership 2 Condominium 3 Cooperative S02 1c	642	CHAR(1)
TYPOWND_		643	CHAR(1)

I. WEIGHTS

VARIABLE	ITEM DESCRIPTION	START POSITION	FORMAT
* BASEWTA	The inverse probability of selection for the CU adjusted for subsampling in the field.(combination of steps 1 and 2 in Section X. C.). NOTE: BASEWTA has been suppressed for some sampled CUs (see Section IV. A.) *D(961) BLS derived	1540	NUM(10,4)
FINLWT21	CU replicate weight # 45 (total sample weight) BLS derived	148	NUM(11,3)
<i>The following are the 44 half sample replicate weights, WTREP01 through WTREP44, which are used for variance computation. They are all BLS derived variables.</i>			
WTREP01	CU replicate weight # 01	687	NUM(11,3)
WTREP02	CU replicate weight # 02	698	NUM(11,3)
WTREP03	CU replicate weight # 03	709	NUM(11,3)
WTREP04	CU replicate weight # 04	720	NUM(11,3)
WTREP05	CU replicate weight # 05	731	NUM(11,3)
WTREP06	CU replicate weight # 06	742	NUM(11,3)
WTREP07	CU replicate weight # 07	753	NUM(11,3)
WTREP08	CU replicate weight # 08	764	NUM(11,3)
WTREP09	CU replicate weight # 09	775	NUM(11,3)
WTREP10	CU replicate weight # 10	786	NUM(11,3)
WTREP11	CU replicate weight # 11	797	NUM(11,3)
WTREP12	CU replicate weight # 12	808	NUM(11,3)
WTREP13	CU replicate weight # 13	819	NUM(11,3)
WTREP14	CU replicate weight # 14	830	NUM(11,3)
WTREP15	CU replicate weight # 15	841	NUM(11,3)
WTREP16	CU replicate weight # 16	852	NUM(11,3)

WTREP17	CU replicate weight # 17	863	NUM(11,3)
WTREP18	CU replicate weight # 18	874	NUM(11,3)
WTREP19	CU replicate weight # 19	885	NUM(11,3)
WTREP20	CU replicate weight # 20	896	NUM(11,3)
WTREP21	CU replicate weight # 21	907	NUM(11,3)
WTREP22	CU replicate weight # 22	918	NUM(11,3)
WTREP23	CU replicate weight # 23	929	NUM(11,3)
WTREP24	CU replicate weight # 24	940	NUM(11,3)
WTREP25	CU replicate weight # 25	951	NUM(11,3)
WTREP26	CU replicate weight # 26	962	NUM(11,3)
WTREP27	CU replicate weight # 27	973	NUM(11,3)
WTREP28	CU replicate weight # 28	984	NUM(11,3)
WTREP29	CU replicate weight # 29	995	NUM(11,3)
WTREP30	CU replicate weight # 30	1006	NUM(11,3)
WTREP31	CU replicate weight # 31	1017	NUM(11,3)
WTREP32	CU replicate weight # 32	1028	NUM(11,3)
WTREP33	CU replicate weight # 33	1039	NUM(11,3)
WTREP34	CU replicate weight # 34	1050	NUM(11,3)
WTREP35	CU replicate weight # 35	1061	NUM(11,3)
WTREP36	CU replicate weight # 36	1072	NUM(11,3)
WTREP37	CU replicate weight # 37	1083	NUM(11,3)
WTREP38	CU replicate weight # 38	1094	NUM(11,3)
WTREP39	CU replicate weight # 39	1105	NUM(11,3)
WTREP40	CU replicate weight # 40	1116	NUM(11,3)
WTREP41	CU replicate weight # 41	1127	NUM(11,3)
WTREP42	CU replicate weight # 42	1138	NUM(11,3)
WTREP43	CU replicate weight # 43	1149	NUM(11,3)
WTREP44	CU replicate weight # 44	1160	NUM(11,3)

m. SUMMARY EXPENDITURE DATA

The variables FOODTOT through HOUSKEEP contain summary expenditure data. They are all BLS derived. The UCC's comprising each summary expenditure variable are listed below the variable description.

VARIABLE	ITEM DESCRIPTION	START POSITION	FORMAT
FOODTOT	Food, total 010110 through 190902, 200112	1171	NUM(12,5)
FOODHOME	Food at home, total 010110 through 180710, 200112	1183	NUM(12,5)
CEREAL	Cereal and cereal products 010110 through 010320	1195	NUM(12,5)
BAKEPROD	Bakery products 020110 through 020820	1207	NUM(12,5)
BEEF	Beef 030110 through 030810	1219	NUM(12,5)
PORK	Pork 040110 through 040610	1231	NUM(12,5)
OTHMEAT	Other meats 050110 through 050900	1243	NUM(12,5)
POULTRY	Poultry 060110 through 060310	1255	NUM(12,5)
SEAFOOD	Fish and seafood 070110 through 070240	1267	NUM(12,5)
EGGS	Eggs 080110	1279	NUM(12,5)
MILKPROD	Fresh milk and cream 090110 through 090210	1291	NUM(12,5)
OTHDAIRY	Other dairy products 100110 through 100510	1303	NUM(12,5)
FRSHFRUT	Fresh fruits 110110 through 110510	1315	NUM(12,5)
FRSHVEG	Fresh vegetables 120110 through 120410	1327	NUM(12,5)
PROCFRUT	Processed fruits 130110 through 130320	1339	NUM(12,5)
PROCVeg	Processed vegetables	1351	NUM(12,5)

	140110 through 140420		
SWEETS	Sugar and other sweets 150110 through 150310	1363	NUM(12,5)
NONALBEV	Nonalcoholic beverages 170110 through 170530, 200112	1375	NUM(12,5)
OILS	Fats and oils 160110 through 160320	1387	NUM(12,5)
MISCFOOD	Miscellaneous foods 180110 through 180710	1399	NUM(12,5)
FOODAWAY	Food away from home 190110 through 190902	1411	NUM(12,5)
ALCBEV	Alcoholic beverages 200111, 200210 through 200530	1423	NUM(12,5)
SMOKSUPP	Tobacco products and smoking supplies 630110 through 630900	1435	NUM(12,5)
PET_FOOD	Pet food 610310	1447	NUM(12,5)
PERSPROD	Personal care products 640110 through 640420	1459	NUM(12,5)
PERSSERV	Personal care services 650110 through 650900	1471	NUM(12,5)
DRUGSUPP	Non-prescription drugs and supplies 550110 through 550900, 570901 through 570902	1483	NUM(12,5)
HOUSKEEP	Housekeeping supplies and services 330110 through 340120	1495	NUM(12,5)

2. MEMBER CHARACTERISTICS AND INCOME FILE (MEMB)

The "MEMB" file, also referred to as the "Member Characteristics and Income" file, contains selected characteristics and earnings for each CU member, including information of relationship to reference person.

Income taxes withheld and pension and retirement contributions are shown both annually and as deductions from the member's last paycheck.

When there is a valid nonresponse, or where nonresponse occurs and there is no imputation, there will be missing values. The type of nonresponse is explained by associated data flag variables described in Section III.C., DATA FLAGS.

a. **CU AND MEMBER IDENTIFIERS**

VARIABLE	ITEM DESCRIPTION	START POSITION	FORMAT
NEWID	CU identification number. Digits 1-7 (CU sequence number, 0000001 through 9999999) uniquely identifies the CU. Digit 8 is the week number, 1 or 2 BLS derived	1	NUM(8)
MEMBNO	Member number S01 1	135	NUM(2)

b. **CHARACTERISTICS OF MEMBERS**

VARIABLE	ITEM DESCRIPTION	START POSITION	FORMAT
CU_CODE1	Member's relationship to reference person CODED 1 Reference person 2 Spouse 3 Child or adopted child 4 Grandchild 5 In-law 6 Brother or sister 7 Mother or father 8 Other related persons 9 Unrelated persons S01 4	70	CHAR(1)
CU_C_DE1		71	CHAR(1)
AGE	Age of member S01 9	9	NUM(2)
AGE_		11	CHAR(1)
RACE	Race of member CODED 1 White 2 Black 3 American Indian, Aleut, Eskimo 4 Asian or Pacific Islander 5 Other S01 10	151	CHAR(1)
RACE_		152	CHAR(1)

SEX	Sex of member CODED 1 Male 2 Female S01 6	174	CHAR(1)
SEX_		175	CHAR(1)
MARITAL	Marital status of member CODED 1 Married 2 Widowed 3 Divorced 4 Separated 5 Never married S01 12	133	CHAR(1)
MARITAL_		134	CHAR(1)
ORIGIN	Origin or ancestry of member CODED 1 European: German Italian Irish French Polish Russian English Scottish Dutch Swedish Hungarian 2 Spanish: Mexican American Chicano Mexican Puerto Rican Cuban Central or South American Other Spanish 3 Afro-American 4 Other and Don't know S01 11	140	CHAR(1)
ORIGIN_		141	CHAR(1)
* EDUCA	Member's highest school grade attended CODED 00 Never attended school 01-11 <i>First grade through eleventh grade</i> 38 <i>Twelfth grade - no degree</i> 39 <i>High school graduate</i>	72	CHAR(2)

40 Some college - no degree
 41 AA degree (occupational/vocational)
 42 AA degree (academic)
 43 Bachelors degree
 44 Masters degree
 45 Professional degree
 46 Doctorate degree

S01 13a

EDUCA_		74	CHAR(1)
* COMPLET	Was highest school grade completed? CODED 1 Yes, grade was completed 2 No, grade was not completed *D(961)	70	CHAR(1)

S01 13b

* COMPLET_	*D(961)	71	CHAR(1)
IN_COLL	Is member currently attending college? CODED 1 Full time 2 Part time 3 Not at all	244	CHAR(1)

S01 13b

IN_COLL_		245	CHAR(1)
ARM_FORC	Is member in the Armed Forces? CODED 1 Yes, member is in the Armed Forces 2 No, member is not in the Armed Forces	242	CHAR(1)

S01 14

ARM__ORC		243	CHAR(1)
SCHLNCHQ	Number of weeks during the past 30 days that school meals were purchased	162	NUM(2)

S02 5b(d)

SCHL_CHQ		164	CHAR(1)
SCHLNCHX	Usual weekly expense for meals purchased at school	165	NUM(8)

S02 5b(c)

SCHL_CHX		173	CHAR(1)
----------	--	-----	---------

c. **WORK EXPERIENCE OF MEMBERS**

VARIABLE	ITEM DESCRIPTION	START POSITION	FORMAT
WKS_WRKD	Number of weeks worked by member full or part time in last 12 months S04A 2	225	NUM(2)
WKS__RKD		227	CHAR(1)
HRSPERWK	Number of hours usually worked per week by member S04A 3	113	NUM(3)
HRSP_RWK		116	CHAR(1)
OCCULIST	The job in which member received the most earnings during the past 12 months fits best in the following category CODED Manager, professional 01 Administrator, manager 02 Teacher 03 Professional Administrative support, technical, sales 04 Administrative support, including clerical 05 Sales, retail 06 Sales, business goods and services 07 Technician Service 08 Protective service 09 Private household service 10 Other service Operator, assembler, laborer 11 Machine operator, assembler, inspector 12 Transportation operator 13 Handler, helper, laborer Precision production, craft, repair 14 Mechanic, repairer, precision production 15 Construction, mining Farming, forestry, fishing 16 Farming 17 Forestry, fishing, groundskeeping Armed forces 18 Armed forces S04A 4a	137	CHAR(2)
OCCU_IST		139	CHAR(1)
EMPLTYPE	Employer from which member received the most earnings in past 12 months CODED 1 Private company, business, or individual	75	CHAR(1)

- 2 Federal government
- 3 State government
- 4 Local government
- 5 Self-employed in own business, professional practice or farm
- 6 Family business or farm, working without pay

S04A 4b

EMPL_YPE		76	CHAR(1)
----------	--	----	---------

WHYNOWRK	Reason member did not work last 12 months CODED	223	CHAR(1)
----------	--	-----	---------

- 1 Retired
- 2 Taking care of home/family
- 3 Going to school
- 4 Ill, disabled, unable to work
- 5 Unable to find work
- 6 Doing something else

S04A 5

WHYN_WRK		224	CHAR(1)
----------	--	-----	---------

d. **INCOME**

VARIABLE	ITEM DESCRIPTION	START POSITION	FORMAT
WAGEX	Amount of wage or salary income before deductions received by member in past 12 months	214	NUM(8)
	S04A 6a		
WAGEX_		222	CHAR(1)
GROSPAYX	Amount of member's last gross pay	95	NUM(8)
	S04A 9		
GROS_AYX		103	CHAR(1)
BSNSX	Amount of income or loss after expenses from own nonfarm business, partnership, or professional practice received by member in past 12 months	61	NUM(8)
	*L		
	S04A 6b		
BSNSX_		69	CHAR(1)
FARMX	Amount of income or loss after expenses from own farm	77	NUM(8)
	*L		

S04A 6c			
FARMX_		85	CHAR(1)
ANYSSINC	Were Social Security checks received by member in the past 12 months?	59	CHAR(1)
	CODED		
	1 Yes		
	2 No		
S04A 7a			
ANYS_INC		60	CHAR(1)
ANYRAIL	Were Railroad Retirement checks received by member in the past 12 months?	57	CHAR(1)
	CODED		
	1 Yes		
	2 No		
S04A 7b			
ANYRAIL_		58	CHAR(1)
SOCRRX	Annual amount of Social Security and Railroad Retirement income received by member in past 12 months	233	NUM(8)
BLS derived			
SOCRRX_		241	CHAR(1)
SS_RRX	Amount of last Social Security or Railroad Retirement payment	183	NUM(8)
S04A 7d			
SS_RRX_		191	CHAR(1)
MEDICARE	Does amount of last Social Security or Railroad Retirement check reflect deduction for Medicare premium?	246	CHAR(1)
	CODED		
	1 Yes		
	2 No		
S04A 7e			
MED_CARE		247	CHAR(1)
SS_RRQ	Number of Social Security or Railroad Retirement checks received by member in past 12 months	228	NUM(4)
S04A 7f			
SS_RRQ_		232	CHAR(1)
US_SUPP	Were Supplemental Security Income checks from the U.S. Government received by member during the past 12	212	CHAR(1)

	months? CODED 1 Yes 2 No		
	S04A 8a		
US_SUPP_		213	CHAR(1)
STA_SUPP	Were any Supplemental Security Income checks from the State or local government received by member during the past 12 months? CODED 1 Yes 2 No	192	CHAR(1)
	S04A 8b		
STA__UPP		193	CHAR(1)
SUPPX	Amount of Supplemental Security income received in the past 12 months	203	NUM(8)
	S04A 8b		
SUPPX_		211	CHAR(1)

e. **TAXES**

VARIABLE	ITEM DESCRIPTION	START POSITION	FORMAT
ANFEDTXX	Annual amount of Federal income tax deduction BLS derived	12	NUM(8)
ANFE_TXX		20	CHAR(1)
FEDTXX	Amount of Federal income tax deducted from last pay S04A 10a	86	NUM(8)
FEDTXX_		94	CHAR(1)
ANSTATXX	Annual amount of state and local income tax deduction BLS derived	48	NUM(8)
ANST_TXX		56	CHAR(1)
STATXX	Amount of state and local income taxes deducted from last pay S04A 10b	194	NUM(8)

STATXX_	202	CHAR(1)
---------	-----	---------

f. RETIREMENT AND PENSION DEDUCTIONS

VARIABLE	ITEM DESCRIPTION	START POSITION	FORMAT
JSSDEDX	Social Security payment during the past 12 months	126	NUM(6)
	BLS derived		
JSSDEDX_		132	CHAR(1)
SLFEMPSS	Amount of contributions to Social Security if member self-employed	176	NUM(6)
	BLS derived		
SLFE_PSS		182	CHAR(1)
ANRRX	Annual amount of Railroad Retirement deduction	39	NUM(8)
	BLS derived		
ANRRX_		47	CHAR(1)
RRX	Amount of Railroad Retirement deducted from last pay	153	NUM(8)
	S04A 10d		
RRX_		161	CHAR(1)
ANGVX	Annual amount of government retirement deduction	21	NUM(8)
	BLS derived		
ANGVX_		29	CHAR(1)
GVX	Amount of government retirement deducted from last pay	104	NUM(8)
	S04A 10e		
GVX_		112	CHAR(1)
ANPVTX	Annual amount of private pension fund deduction	30	NUM(8)
	BLS derived		
ANPVTX_		38	CHAR(1)
PVTX	Amount of private pension fund deducted from last pay	142	NUM(8)
	S04A 10f		

PVTX_		150	CHAR(1)
IRAX	Amount of money placed in an individual retirement plan, such as IRA or Keogh, in past 12 months	117	NUM(8)
	S04A 13b		
IRAX_		125	CHAR(1)

3. **DETAILED EXPENDITURES (EXPN) FILE**

The "EXPN" file, also referred to as the "Detailed Expenditures" file, provides weekly expenditure data at the most detailed level available. In this file, each expenditure reported by a CU is identified by a Universal Classification Code (UCC), NEWID, gift/nongift status, and day on which the expenditure occurred. UCC's are six digit codes that identify items or groups of items. (See Appendix 2.A for a listing of UCC's.) There may be more than one record for a UCC on a single day if that is what was reported in the diary. There are no missing values in this file. If a value was not reported then there is not a record for the UCC.

VARIABLE	ITEM DESCRIPTION	START POSITION	FORMAT
NEWID	CU identification number. Digits 1-7 (CU sequence number, 0000001 through 9999999) uniquely identifies the CU. Digit 8 is the week number, 1 or 2	1	NUM(8)
ALLOC	Adjustment status for cost variable CODED 0 Not allocated or topcoded 1 Allocated, not topcoded 2 Topcoded and allocated 3 Topcoded, not allocated	9	CHAR(1)
COST	Total cost of item, including sales tax	10	NUM(12,5)
GIFT	Was item bought for someone outside the CU? CODED 1 Yes 2 No	22	CHAR(1)
PUB_FLAG	Is cost included in published bulletin? CODED 1 Not published 2 Published in Integrated bulletin	23	CHAR(1)
*QREDATE	Purchase date recode field Consists of: Sequential day of the Diary week (1-7) Day of the week, Sunday through Saturday (1-7) Reference month of this expenditure, (01-12) Reference day of this expenditure, (01-31) Reference year of this expenditure, (00-99)	24	CHAR(10)

QREDATE_		34	CHAR(1)
UCC	Universal Classification Code See Appendix 2.A-- EXPENDITURE UCC'S ON EXPN FILE	35	CHAR(6)

4. **INCOME (DTAB) FILE**

The "DTAB" file, also referred to as the "Income" file, contains CU characteristic and income data. This file is created directly from the FMLY file. It was created to facilitate computer processing when linking CU income and demographic characteristics data with EXPN expenditure data. As such, the file structure is similar to EXPN. Each demographic characteristic and income item is identified by UCC (See Appendix 2.B for a listing of UCC's). There are no records with missing values in DTAB. If the corresponding FMLY file variable contained a missing value, there is no record for the UCC.

VARIABLE	ITEM DESCRIPTION	START POSITION	FORMAT
NEWID	CU identification number. Digits 1-7 (CU sequence number, 0000001 through 9999999) uniquely identifies the CU. Digit 8 is the week number, 1 or 2	1	NUM(8)
UCC	Universal Classification Code See Appendix 2.B--INCOME AND RELATED UCC'S ON DTAB FILE	9	CHAR(6)
AMOUNT	Amount of UCC	15	NUM(12)
AMOUNT_	CODED T – Topcoded Blank -- Not topcoded	27	CHAR(1)
*PUB_FLAG	Is amount included in published bulletin? CODED 1 Not published 2 Published in Integrated bulletin	28	CHAR(1)

5. **PROCESSING FILES**

a. **AGGregation file**

DIARY96\AGGD96.TXT

The AGG file shows which UCC's go into each category listed in the sample table produced by the microdata file verification and estimation program. (See Section VII). It designates each category with a unique 6-digit line number. It is formatted as follows:

DESCRIPTION	START POSITION	FORMAT
UCC (Universal Classification Code)	3	CHAR(6)
Line Number: represents a line in the sample table	15	CHAR(6)

b. LABel file

DIARY96\LABELD96.TXT

The LABEL file assigns an identification label to each AGG file line number. It is formatted as follows:

DESCRIPTION	START POSITION	FORMAT
Line Number: represents a line in the sample table	1	CHAR(6)
Label: descriptive label in the sample table (with leading blanks)	10	CHAR(48)

c. UCC file

DIARY96\UCCD96.TXT

The UCC file contains UCC's and their abbreviated titles, identifying the expenditure, income, or demographic item represented by each UCC. It is formatted as follows:

DESCRIPTION	START POSITION	FORMAT
Universal Classification Code (UCC) (See Appendix 2 for a list of UCC's and their full titles by file--Expenditure (EXPEN) or income (DTAB))	1	CHAR(6)
UCC title	8	CHAR(50)

d. SAMPLe program file

DIARY96\ SAMPLD96.TXT

The SAMPLD96 file contains the computer program used in Section VII.A. of the documentation. This file has been created to provide programming assistance.

IV. TOPCODING AND OTHER NONDISCLOSURE REQUIREMENTS

Sensitive CU data are changed so that users will not be able to identify CUs who participated in the survey. Topcoding refers to the replacement of data in cases where the value of the original data exceeds prescribed critical values. Critical values for each variable containing sensitive data are chosen in accordance with Census Disclosure Review Board guidelines. Each observation that falls outside the critical value is replaced with a topcoded value that represents the mean of the subset of all outlying observations. All four quarters of data in the CE microdata release are used when calculating the topcode amount. If an observation is topcoded, the flag variable assigned to that variable is set to 'T', indicating that topcoding has occurred.

Since the mean of the set of values outside the critical value may differ with each annual (four-quarter) release, the topcode values may change annually. By changing the topcode values in accordance with the observations above the critical values, the first moment will be preserved for each four-quarter data release when using the total sample. This, however, will not be the case when means are estimated by characteristic, because topcode values are not calculated by characteristic.

Some variables will have a critical value, but no topcode amount. This implies that there are no observations outside the critical value on the current five-quarter release.

A. CU CHARACTERISTICS AND INCOME FILE (FMLY)

The following FMLY file variables are subject to topcoding.

AGE_REF	Age of reference person
AGE2	Age of spouse
ADDFEDX	Amount of Federal income tax paid in addition to that withheld
ADDOTHX	Amount of other taxes paid but not reported elsewhere
ADDSTAX	Amount of state and local income tax paid in addition to that withheld
ALIOTHX	Amount received from regular contributions by all CU members
CHDLMPX	Amount received by all CU members for a lump sum child support payment in last 12 months
CHDOTHX	Amount received by all CU members in last 12 months for other child support
DIVX	Amount received from dividends, royalties, estates, or trusts
FEDREFX	Amount of refund from Federal income tax
INSREFX	Amount of refund from insurance policies
INTX	Amount received from interest on savings accounts, or bonds
LUMPX	Amount from lump sum payments from estates, trusts, royalties, alimony, child support, prizes, games of chance, or persons outside CU
OCCEXPX	Amount paid by CU for occupational expenses, last 12 months
OTHINX	Amount from other money income, including money from care of foster children, cash scholarships and fellowships, or stipends, not based on working
OTHREFX	Amount of refund from other sources, including any other taxes
OTHRNTX	Amount of net income or loss received from other rental units
PENSIONX	Amount received from pensions or annuities from private companies, military or government, IRA or Keogh
PTAXREFX	Amount of refund from property taxes
ROOMX	Amount of net income or loss received from roomers or boarders
SALEX	Amount received from sale of household furnishings, equipment, clothing, jewelry, pets or other belongings, excluding sale of vehicles or property
SSREFX	Amount of refund from overpayment on Social Security
STATREFX	Amount of refund from state or local income tax
TAXPROPX	Amount of personal property taxes paid but not reported elsewhere

The critical values and topcode values associated with the above variables follow.

<u>Variable</u>	<u>Critical value +</u>	<u>Critical value -</u>	<u>Topcode value +</u>	<u>Topcode value -</u>
ADDFEDX	16,000	-	38,656	-
ADDOTHX	3,500	-	5,706	-
ADDSTAX	3,000	-	18,686	-
AGE_REF	90	-	93	-
AGE2	90	-	91	-
ALIOTHX	24,000	-	34,273	-
CHDLMPX	6,600	-	10,011	-
CHDOTHX	14,400	-	19,000	-
DIVX	30,000	-	110,178	-
FEDREFX	4,500	-	9,508	-
INSREFX	30,000	-	100,000	-
INTX	35,000	-	282,333	-
LUMPX	75,000	-	272,308	-
OCCEXPX	2,000	-	5,639	-
OTHINX	25,000	-	40,000	-
OTHREFX	2,100	-	7,440	-
OTHRNTX	26,800	-17,000	41,217	-29,632
PENSIONX	45,000	-	61,300	-
PTAXREFX	1,200	-	2,170	-
ROOMX	9,600	-7,000	37,921	-28,000
SALEX	5,000	-	16,857	-
SSREFX	6,500	-	-	-
STATREFX	1,300	-	3,122	-
TAXPROPX	4,000	-	7,404	-

Some income variables are constructed by summing up the values of “lower level” MEMB or FMLY file component variables. These types of variables that are subject to topcoding are not topcoded in the conventional sense of being replaced by a topcode value, but rather are flagged as topcoded because one of their component variables is topcoded. Examples of situations that may occur are illustrated below.

The value for the variable FBSNSX (family income from non-farm business) is computed as the sum of the values reported for the variable BSNSX (member income from non-farm business) from the MEMB file. BSNSX is subject to topcoding beyond the critical value of \$150,000 (-\$40,000). The topcode value for BSNSX is \$237,633 (-\$200,000).

BSNSX			FBSNSX		
<u>CU</u>		<u>REPORTED</u>	<u>AFTER TOPCODING</u>	<u>VALUE</u>	<u>FLAGGED AS TOPCODED?</u>
CU 1:	MEMB1	\$150,000	\$150,000		
	MEMB2	150,000	150,000	300,000	No
CU 2:	MEMB1	282,000	237,633		
	MEMB2	18,000	18,000	255,633	Yes
CU 3	MEMB1	160,000	237,633		
	MEMB2	60,000	60,000	297,633	Yes
CU 4	MEMB1	140,000	140,000		
	MEMB2	-50,000	-200,000	-60,000	Yes

While CUs 1 and 2 each originally report \$300,000 in BSNSX, topcoding is done only on the value reported by MEMB1 of CU2. Thus, the value for FBSNSX for CU2 is lower than for CU1 and is flagged as topcoded while CU1 is not. By using the mean of the subset of observations that are above (below) the critical value as the topcode amount, values on the public use data can be either below or above the actual reported value. Note that while CU2 has a topcoded value below the reported value, CU3's topcoded FBSNSX value (\$297,633) is higher than the amount that it reported (\$220,000). The case of CU4 demonstrates that the reported value for FBSNSX can be positive, while the topcoded value can be negative. This is due to a topcoded negative BSNSX value for MEMB2 that is large enough to change total CU income from positive to negative.

Components of income are used to compute income values at the consumer unit level. If an income component (MEMB or FMLY level) is topcoded, the topcoded value is used in the family level calculation. (See the descriptions of each variable in Sections III.E.1.e. - III.E.1.h. for a list of component variables.) The following are the income variables which are calculated using values of their component variables:

EARNX	Amount of CU income from earnings before taxes
FBSNSX	Amount of income from non-farm business
FFARMX	Amount of income or loss received from own farm
FFEDTXX	Federal tax deducted from last pay, annualized for all CU members
FGVX	Government retirement deducted from last pay, annualized for all CU members
FINCAFTX	Amount of CU income after taxes
FINCBEFX	Amount of CU income before taxes
FIRAX	Amount of money placed in individual retirement plan
FJSSDEDX	Amount of income contributed to Social Security
FPVTX	Private pensions deducted from last pay, annualized for all CU members
FRRX	Railroad retirement deducted from last pay, annualized for all CU members
FSTATXX	State and local income taxes deducted from last pay, annualized for all CU members
FWAGEX	Amount of wage and salary income before deduction
NONERNX	Amount of income other than earnings before taxes
OTHRECX	Amount of other money receipts excluded from family income
PERSTAX	Amount of personal taxes

The value of the state identifying variable, STATE, must be suppressed for some observations in order to meet the Census Disclosure Review Board's criterion for releasing geographic information. This criterion states that the smallest geographically identifiable area must have a population of at least 100,000. STATE data were evaluated vis-a-vis POPSIZE, REGION, and BLS_URBN, which show, respectively, the population size of the geographic area that is sampled, the four Census regions, and whether the area is urban or rural. Thus, some STATE codes were suppressed because, in combination with these variables, they could be used to identify areas of 100,000 or less. Approximately 17 percent

of the STATE variable records on the FMLY files are blank. The STATE flag (STATE_) is given a value of 'T' if STATE is suppressed.

A small proportion of STATE codes are replaced with codes of states other than the state where the CU resides. By re-coding in this manner, suppression of POPSIZE and REGION may be avoided. (In past releases selected observations of POPSIZE and REGION also required suppression.) If an observation of a CU's state of residence is re-coded with another state's code, the flag variable (STATE_) of the re-coded state is assigned an 'R'. The flag variable is also assigned an 'R' for either all or a portion of other observations from that state. In total, approximately 4% of observations of STATE_ are assigned an 'R'.

01	Alabama	*28	Mississippi
02	Alaska	**29	Missouri
RR 04	Arizona	31	Nebraska
*05	Arkansas	R 32	Nevada
**06	California	R 33	New Hampshire
08	Colorado	34	New Jersey
09	Connecticut	*35	New Mexico
10	Delaware	RR**36	New York
R 11	District of Columbia	**37	North Carolina
**12	Florida	RR 39	Ohio
**13	Georgia	**40	Oklahoma
15	Hawaii	**41	Oregon
16	Idaho	42	Pennsylvania
**17	Illinois	45	South Carolina
RR**18	Indiana	*46	South Dakota
*19	Iowa	**47	Tennessee
**20	Kansas	48	Texas
21	Kentucky	49	Utah
22	Louisiana	50	Vermont
R*23	Maine	**51	Virginia
24	Maryland	**53	Washington
25	Massachusetts	R 54	West Virginia
**26	Michigan	55	Wisconsin
**27	Minnesota		

* indicates that the STATE code has been suppressed for all sampled CUs in that state (STATE_ = 'T' for all observations).

** indicates that the STATE code has been suppressed for some sampled CUs in that state (STATE_ = 'T' for some observations).

R indicates that either all observations from this state have been re-coded or all strata¹ of observations from this state include "re-codes" from other states.

RR indicates that either some observations from this state have been re-coded or at least one stratum¹ of observations from this state includes "re-codes" from other states.

R* indicates that the STATE code has been suppressed and, either STATE has been re-coded or the state includes "re-codes" from other states in all strata¹.

RR** indicates that the STATE code has been suppressed and, either STATE has been re-coded or the state includes "re-codes" from other states in at least one stratum¹.

¹ A STATE stratum is a unique POPSIZE and BLS_URBN combination.

States not listed are not in the CE sample.

B. MEMBER CHARACTERISTICS AND INCOME FILE (MEMB)

The following MEMB file variables are subject to topcoding.

AGE	Age of member
ANFEDTXX	Annual amount of Federal income tax deduction
ANGVX	Annual amount of government retirement deduction
ANPVTX	Annual amount of private pension fund deduction
ANRRX	Annual amount of Railroad Retirement deduction
ANSTATXX	Annual amount of state and local income tax deduction
BSNSX	Amount of income or loss from nonfarm business, partnership, or professional practice
FARMX	Amount of income or loss after expenses from own farm
FEDTXX	Amount of Federal income tax deduction from last pay
GROSPAYX	Amount of last gross pay
GVX	Amount of government retirement deduction from last pay
IRAX	Amount of money placed in an individual retirement plan, such as IRA or Keogh
JSSDEDX	Amount of Social Security payment
PVTX	Amount of private pension fund deduction from last pay
RRX	Amount of Railroad Retirement deduction from last pay
SLFEMPSS	Amount of self employment Social Security contributions
STATXX	Amount of state and local income tax deduction from last pay
WAGEX	Amount of wage and salary income before deductions

The critical values and topcode values associated with the above variables follow.

<u>Variable</u>	<u>Critical value +</u>	<u>Critical value -</u>	<u>Topcode value +</u>	<u>Topcode value -</u>
AGE	90	-	93	-
ANFEDTXX	15,235	-	29,126	-
ANGVX	5,970	-	7,392	-
ANPVTX	10,000	-	15,108	-
ANRRX	15,294	-	-	-
ANSTATXX	4,988	-	8,759	-
BSNSX	150,000	-40,000	237,633	-200,000
FARMX	40,000	-40,000	200,000	-
FEDTXX	777	-	1,721	-
GROSPAYX	4,308	-	12,746	-
GVX	435	-	1,320	-
IRAX	12,000	-	30,909	-
JSSDEDX	5,450	-	7,638	-
PVTX	520	-	881	-
RRX	600	-	-	-
SLFEMPSS	11,000	-	13,249	-
STATXX	235	-	551	-
WAGEX	150,000	-	250,628	-

Special suppression for MEMB file variables

The five MEMB file variables--FEDTXX, GVX, PVTX, RRX, and STATXX--describe deductions from the most recent pay. These variables are used in conjunction with GROSPAYX (amount of last gross pay) and WAGEX (annual wage and salary income) to derive ANFEDTXX, ANGVX, ANPVTX, ANRRX, and ANSTATXX, which represent the estimated annual deductions for each of these income deduction categories. For example, the estimated annual Federal income tax deduction from pay is calculated as

$$(1) \quad \text{ANFEDTXX} = (\text{WAGEX} (\text{FEDTXX}/\text{GROSPAYX})).$$

Note that WAGEX can be estimated by using the above terms and rearranging such that

$$(2) \quad \text{WAGEX} = (\text{ANFEDTXX} (\text{GROSPAYX}/\text{FEDTXX})).$$

In this case, a problem with confidentiality may arise when neither ANFEDTXX, GROSPAYX, nor FEDTXX (calculation components) are beyond the critical value for topcoding, *but WAGEX is*. In this situation WAGEX is topcoded on the public use microdata file, however it can be recalculated in its non-topcoded form by using the calculation components of (2) above. In order to prevent this, the calculation components of equation (2) will be suppressed (blanked out) and their associated flags will be given a 'T', *if WAGEX is greater than the critical value but all its calculation variables are below their critical values*. The following chart describes in detail the specific rules that will be applied to prevent the potential breach of confidentiality outlined above.

If WAGEX is greater than the critical value but ANFEDTXX, GROSPAYX, and FEDTXX are not, then suppress the values for ANFEDTXX, GROSPAYX, and FEDTXX and give their flag variables a value of 'T'.

If WAGEX is greater than the critical value but ANGVX, GROSPAYX, and GVX are not, then suppress the values for ANGVX, GROSPAYX, and GVX and give their flag variables a value of 'T'.

If WAGEX is greater than the critical value but ANPVTX, GROSPAYX, and PVTX are not, then suppress the values for ANPVTX, GROSPAYX, and PVTX and give their flag variables a value of 'T'.

If WAGEX is greater than the critical value but ANRRX, GROSPAYX, and RRX are not, then suppress the values for ANRRX, GROSPAYX, and RRX and give their flag variables a value of 'T'.

If WAGEX is greater than the critical value but ANSTATXX, GROSPAYX, and STATXX are not, then suppress the values for ANSTATXX, GROSPAYX, and STATXX and give their flag variables a value of 'T'.

C. DETAILED EXPENDITURE FILE (EXPX)

The EXPX variable COST is subject to topcoding for the following UCCs (Universal Classification Codes).

UCC	Description
001000	Purchase price of stocks, bonds, mutual funds
210110	Rent of dwelling, includes parking fees
210210	Lodging away from home
210310	Housing for someone at school
210900	Ground or land rent
550320	Medical equipment for general use
550330	Supportive convalescent and medical equipment
560110	Physicians' services (hospital care)
560210	Dental services
560310	Eye care services, including eye examination or surgery and combined eye care services
560320	Services by practitioner other than physicians
560330	Lab tests and x-rays
560900	Nursing services and therapeutic treatments
570000	Hospital care not specified
570220	Care in convalescent or nursing home

570230 Other medical care service
570901 Rental of medical or surgical equipment

If the value of COST is greater (less) than the designated critical values for the above UCC's, COST is set to the topcode value and the associated flag variable, COST_, is set to 'T'. The critical values and topcode values (rounded to the nearest dollar) of the variable COST that are associated with the above UCCs follow.

<u>UCC</u>	<u>Critical value +</u>	<u>Critical value -</u>	<u>Topcode value +</u>	<u>Topcode value -</u>
001000	92,000	-	-	-
210110	1,000	-	1,456	-
210210	1,000	-	-	-
210310	2,500	-	-	-
210900	1,333	-	-	-
550320	300	-	-	-
550330	675	-	2,459	-
560110	600	-	7,234	-
560210	1,250	-	1,888	-
560310	300	-	-	-
560320	900	-	-	-
560330	575	-	1,250	-
560900	1,000	-	-	-
570000	1,850	-	8,000	-
570220	10,300	-	-	-
570230	1,000	-	-	-
570901	500	-	832	-

D. INCOME FILE (DTAB)

The DTAB variable AMOUNT is subject to topcoding for the following UCCs (Universal Classification Codes).

<u>UCC</u>	<u>Description</u>
800910	Deductions for government retirement
800920	Deductions for Railroad Retirement
800931	Deductions for private pensions
800940	Deductions for Social Security
900000	Wages and salaries
900010	Net business income
900020	Net farm income
900030	Social Security and Railroad Retirement income
900040	Pensions and annuities
900050	Dividend/royalty/estate/trust
900060	Roomer and boarder income
900070	Other rental income
900080	Interest
900090	Supplemental security income
900100	Unemployment compensation
900110	Workers' compensation
900120	Public assistance
900131	Child support payments received
900132	other regular contributions received, including alimony

<u>UCC</u>	<u>Description</u>
900140	Other income
900150	Food stamps
910000	Lump sum payments from estates, trusts, etc.
910010	Money from sale household furnishings etc.
910020	Overpayment on Social Security
910030	Refund from insurance policies
910040	Refunds from property taxes
910041	Lump sum child support payments received
950000	Federal income tax
950001	Federal income tax refunds
950010	State/local income tax
950011	State/local income tax refunds
950021	Other taxes
950022	Personal property taxes
950023	Other tax refunds
980000	Income before taxes
980070	Income after taxes

Note: Data in the DTAB file are selected data from the FMLY file expressed in a file structure similar to the EXPN structure. This facilitates processing.

AMOUNT for the following UCC's is topcoded because the FMLY file variables corresponding to these UCC's are topcoded due to recalculation. (See Section IV.A. on topcoding of FMLY variables.)

<u>UCC</u>	<u>FMLY variable</u>	<u>Description</u>
800910	FGVX	Deductions for government retirement
800920	FRRX	Deductions for rail road retirement
800931	FPVTX	Deductions for private pensions
800932	FIRAX	Self-employment retirement plan contributions
800940	FJSSDEDX	Deductions for social security
900000	FWAGEX	Wages and salaries
900010	FBSNSX	Net business income
900020	FFARMX	Net farm income
980000	FINCBEFX	Income before taxes
980070	FINCAFTX	Income after taxes

V. ESTIMATION PROCEDURES

This section provides users of the CE Diary microdata files with procedures for estimating means and variances of data associated with any U.S. subpopulation. The production of *Consumer Expenditures in 1996, Report 926* (1998) used an integration methodology which incorporated information from both Diary and Interview Surveys. In addition, users will not be able to match all values because of suppression of some values, due to topcoding. See the topcoding and other nondisclosure requirements in Section IV.

A. DEFINITION OF TERMS

Consider the following general situation. We wish to estimate expenditures on certain food items for a special group (subpopulation) of U.S. CUs; for example, all CUs of three persons. Our specific objective is to estimate the expenditures for item k over a period of q months, where data collected over r months are used in the estimate. The following definitions will be helpful in formulating the above type of estimate.

Definition of Terms:

Let

- S = all CUs in the subpopulation of interest
- x = expenditure item(s) of interest
- q = number of months for which estimate is desired
- r = number of months in which expenditures were made to be used in calculating the estimate
- D = number of days in each of the months in which expenditures were made
- j = individual CU in subpopulation S
- t = month of expenditure

Then

- $X_{(j,k,t)}$ = the amount of money CU (j) spent on item k for a week during month t
- $W_{(j,t,F21)}$ = the weight assigned to CU (j) during month t

The F21 denotes FINLWT21 which is used for population estimates.

NOTE: The CUs on the Diary Survey microdata files represent the U.S. population. Some CUs represent more of the population than others; and hence carry more weight. The weight, $W_{(j,t,F21)}$, is a complex estimate of this representation. Refer to Section X.C. WEIGHTING for an explanation of weights. The weights have been adjusted so that the sum of all CU weights for one month approximates one third of the U.S. population. Consequently, the weights for three months (one quarter) of data approximate the total U.S. population.

Using the above terminology, we may define:

- $X_{(S,k)(q,r)}$ as an estimate for the expenditures of subpopulation S on item k over a period of q months, where data collected over r months are used.

and

- $\bar{X}_{(S,k)(q,r)}$ as an estimate of the mean expenditures of subpopulation S on item k over a period of q months, where data collected over r months are used.

B. ESTIMATION OF TOTAL AND MEAN EXPENDITURES

As an example, let us estimate total expenditures on milk (item k) of subpopulation S over a 12-month period. Data collected over 6 months will be used to make the estimate. Users may use less than

12 months of data to perform seasonal calculations. In the notation described above, the estimate is $X_{(S,k)(12,6)}$.

$$X_{(S,k)(12,6)} = 3 \left(\frac{12}{6} \right) \sum_{t=1}^6 \left(\sum_{j=1}^n \left(\frac{D_{(t)}}{7} \right) W_{(j,t,F21)} X_{(j,k,t)} \right)_t \quad (1a)$$

where the inner summation sums expenditures for all j in S , indexed from $j = 1$ through n and the outer summation sums over months $t = 1$ through 6. The factor "3" compensates for the fact that the weights for the CUs visited in one month have been adjusted to represent one third of the U.S. population. The factor "12" reflects our desire to estimate expenditures over a 12-month period; and the "6" is the adjustment made because data for 6 months are used. Since the data $X_{(j,k,t)}$ are in terms of weekly expenditures, the factors, (number of days in the month)/7, are used to convert weekly expenditures into their monthly equivalents.

The above formula can be generalized to estimate the total expenditures of subpopulation S on item k for q months, but using data collected over r months. The generalization is

$$X_{(S,k)(q,r)} = 3 \left(\frac{q}{r} \right) \sum_{t=1}^r \left(\sum_{j=1}^n \left(\frac{D_{(t)}}{7} \right) W_{(j,t,F21)} X_{(j,k,t)} \right)_t \quad (1b)$$

where the inner summation sums expenditures for all j in S , indexed from $j = 1$ through n and the outer summation sums over months $t = 1$ through r .

An estimate for the expenditures for two or more items may be obtained by summing those expenditures at the CU level and then proceeding as before.

The next example will give an estimate, $\bar{X}_{(S,k)(12,6)}$, of mean expenditures over twelve months (q), on item k , of CUs in subpopulation S , where data collected over a six month period (r) are used. The result is

$$\bar{X}_{(S,k)(12,6)} = \frac{3 \left(\frac{12}{6} \right) \sum_{t=1}^6 \left(\sum_{j=1}^n \left(\frac{D_{(t)}}{7} \right) W_{(j,t,F21)} X_{(j,k,t)} \right)_t}{\frac{3 \sum_{t=1}^6 \left(\sum_{j=1}^n W_{(j,t,F21)} \right)_t}{6}} \quad (2a)$$

where the numerator is an estimate of aggregate expenditures as formulated in equation (1a), and where the denominator is an estimate of the population of CUs in the U.S. during the six-month period for which the expenditure data are collected. The inner summation in the denominator of (2a) sums FINLWT21 for a given month (t), for all j in S , indexed from $j = 1$ through n , and the outer summation in the denominator of (2a) sums over months $t = 1$ through 6. As in the estimate of aggregate expenditures, the factor "3" to the left of the outer summation in the denominator of equation (2a) adjusts FINLWT21 to represent the entire population for each month of data used. The proper U.S. population count is arrived at by dividing the denominator by r , or in this case "6", (representing the 6 month period of collected data in this example).

The above formula generalizes to $\bar{X}_{(S,k)(q,r)}$, (i.e., the estimate of the mean expenditure by subpopulation S on item k for q months using data collected over r months). In detail:

$$\bar{X}_{(S,k)(q,r)} = \frac{q \sum_{t=1}^r \left(\sum_{j=1}^n \left(\frac{D_{(t)}}{7} \right) W_{(j,t,F21)} X_{(j,k,t)} \right)}{\sum_{t=1}^r \left(\sum_{j=1}^n W_{(j,t,F21)} \right)} \quad (2b)$$

Note: The factors “3” (adjustment of FINLWT21 to one U.S. population) and “6”, (number of months, r , for which the data are collected), which appear both in the numerator and the denominator of (2a), cancel. These scalars are dropped from the general form of $\bar{X}_{(S,k)(q,r)}$.

The estimates for total ($X_{(S,k)(q,r)}$) and mean expenditures ($\bar{X}_{(S,k)(q,r)}$) are based on all CUs; not just the CUs with positive expenditures for item k . Consider the calculation for the mean expenditure of tobacco. The formula $\bar{X}_{(S,k)(q,r)}$ includes all CUs, both smoking and nonsmoking. One might be more interested in the mean expenditures on tobacco but only for those CUs that actually have expenditures. This can be accounted for by properly defining the initial subpopulation S so as to restrict it to CUs with positive tobacco expenditures.

C. ESTIMATION OF MEAN ANNUAL INCOME

Let $\bar{Z}_{(S,r)}$ be an estimate of the mean annual income of CUs in subpopulation S , where income data collected over r months is to be used.

Let $Z_{(j,t)}$ = the annual income reported by CU_(j) in month t . Then the estimated mean annual income is

$$\bar{Z}_{(S,r)} = \frac{\sum_{t=1}^r \left(\sum_{j=1}^n W_{(j,t,F21)} Z_{(j,t)} \right)}{\sum_{t=1}^r \left(\sum_{j=1}^n W_{(j,t,F21)} \right)}$$

VI. RELIABILITY STATEMENT

A. DESCRIPTION OF SAMPLING ERROR AND NONSAMPLING ERROR

Sample surveys are subject to two types of errors, sampling and nonsampling. Sampling errors occur because observations are not taken from the entire population. The standard error, which is the accepted measure for sampling error, is an estimate of the difference between the sample data and the data that would have been obtained from a complete census. The sample estimate and its estimated standard error enables one to construct confidence intervals.

Assuming the Normal Distribution applies to the means of expenditures, the following statements can be made:

- (1) The chances that an estimate from a given sample would differ from a complete census figure by less than one standard error are approximately 68 out of 100.
- (2) The chances that the difference would be less than 1.6 times the standard error are approximately 90 out of 100.
- (3) The chances that the difference would be less than two times the standard error are approximately 95 out of 100.

Nonsampling errors can be attributed to many sources, such as definitional difficulties, differences in the interpretation of questions, inability or unwillingness of the respondent to provide correct information, mistakes in recording or coding the data obtained, and other errors of collection, response, processing, coverage, and estimation for missing data. The full extent of the nonsampling error is unknown. Estimates using a small number of observations are less reliable. A small amount of nonsampling error can cause a small difference to appear significant even when it is not. It is probable that the levels of estimated expenditure obtained in the Diary Survey are generally lower than the "true" level due to the above factors.

B. ESTIMATING SAMPLING ERROR

Variance estimates for 1996 are not available at this time. This section uses 1985 mean and variance estimates for the examples. Users may generate their own 1996 variance estimates by using the methodology presented in the variance estimation section below.

1. VARIANCE ESTIMATION

Variance estimation can be done in many ways. The method illustrated below (a pseudo-replication technique) is chosen because it is accurate yet simple to understand. The basic idea is to artificially construct several "subsamples" from the original sample data. This construction is done in a manner so that the variance information of the original data is preserved in these subsamples. These subsamples (or pseudo-replications) can then be used to obtain approximate variances for the estimates. Justifications for the procedure may be found below in references 1 and 2.

The Diary microdata files contain information that facilitates this form of variance estimation procedure. Specifically, 45 weights are associated with each CU. The forty-fifth weight, called FINLWT21 at BLS, (which is the weight for the total sample) is used for estimations of total or mean expenditures. The other weights (replicates 1 through 44) are used for variance estimation of the totals or means. Note that half of the weights in each replicate are zero. This reflects the fact that in this technique only half the CUs are used in each of the 44 pseudo-replicates. Recall that $X_{(S,k)(q,r)}$ is an estimate for the expenditures of subpopulation S on item k over a period of q months, where data collected over r months are used. This notation does not reveal the fact that 45 replicate weights are to be used for estimation of variance. We expand the notation to include this information. Specifically, let

$X_{(S,k)(q,r),a}$ = an estimate of the same quantity as $X_{(S,k)(q,r)}$, but using the weights of the a^{th} replicate.

That is $X_{(S,k)(q,r),a}$ is an estimate of the total expenditures by CUs in subpopulation S on item k over q months using r months of collection data, and where the weights from the a^{th} replicate are used. Note that the estimate using any one of the first 44 replicate weights only uses part of the data; hence in general $X_{(S,k)(q,r),a}$ is not equal to $X_{(S,k)(q,r)}$.

An estimate for the variance of $X_{(S,k)(q,r)}$ (denoted by $V(X_{(S,k)(q,r)})$) can be calculated using the following formula:

$$V(X_{(S,k)(q,r)}) = \frac{1}{44} \sum_{a=1}^{44} (X_{(S,k)(q,r),a} - X_{(S,k)(q,r)})^2$$

Estimates for the variances of $\bar{X}_{(S,k)(q,r)}$ and $\bar{Z}_{(S,r)}$ are similar and are given below.

$$V(\bar{X}_{(S,k)(q,r)}) = \frac{1}{44} \sum_{a=1}^{44} (\bar{X}_{(S,k)(q,r),a} - \bar{X}_{(S,k)(q,r)})^2$$

and

$$V(\bar{Z}_{(S,r)}) = \frac{1}{44} \sum_{a=1}^{44} (\bar{Z}_{(S,r),a} - \bar{Z}_{(S,r)})^2$$

where $\bar{X}_{(S,k)(q,r),a}$ and $\bar{Z}_{(S,r),a}$ are estimates similar to $\bar{X}_{(S,k)(q,r)}$ and $\bar{Z}_{(S,r)}$ except weights of the a^{th} replicates are used.

2. STANDARD ERROR OF THE MEAN

The standard error of the mean, $S.E.(\bar{x})$, is defined as the square root of the variance of the mean. $S.E.(\bar{x})$ is used to obtain confidence intervals that evaluate how close the estimate may be to the true population mean. For example, the average weekly expenditures for meats, poultry, fish, and eggs for total complete income reporters in 1985 was \$11.13. The standard error for this estimate is \$.23. Hence, the 95 percent confidence interval around this estimate is from \$10.67 to \$11.59. Therefore, we could conclude with 95 percent confidence that the mean weekly expenditures for meats, poultry, fish, and eggs for total complete income reporters in 1985 lies within the interval \$10.67 to \$11.59.

3. STANDARD ERROR OF THE DIFFERENCE BETWEEN TWO MEANS

Standard errors may also be used to perform hypothesis testing, a procedure for distinguishing between population parameters using sample estimates. The most common types of hypotheses are: 1) the population parameters are identical, versus 2) they are different.

For example, in 1985 the average weekly expenditures for food for complete income reporters in the second income quintile was \$44.17 and for complete income reporters in the third income quintile was \$55.83. The apparent difference between the two mean expenditures is \$55.83 - \$44.17 = \$11.66. The standard error on the estimate of \$44.17 is \$1.36 and the estimated standard error for \$55.83 is \$1.55. The standard error (S.E.) of a difference is approximately equal to

$$S.E.(\bar{X}_1, \bar{X}_2) = \sqrt{V(\bar{X}_1) + V(\bar{X}_2)}$$

where

$$V(\bar{X}_i) = (S.E.(\bar{X}_i))^2$$

This assumes that \bar{X}_1 and \bar{X}_2 are disjoint subsets of the population. Hence, the standard error of the difference in food expenditures between the second and third income quintile groups of complete income reporters is about

$$\sqrt{(1.55)^2 + (1.36)^2} = 2.06$$

This means that the 95 percent confidence interval around the difference is from \$7.54 to \$15.78. Since this interval does not include zero, we can conclude with 95 percent confidence that the mean weekly food expenditures for the third income quintile group of complete income reporters is greater than the mean weekly food expenditures for the second income quintile group.

Analyses of the difference between two estimates can also be performed on nondisjoint sets of population, where one is a subset of the other. The formula for computing the standard error (S.E.) of the difference between two nondisjoint estimates is

$$S.E.(\bar{X}_1, \bar{X}_2) = \sqrt{V(\bar{X}_1) + V(\bar{X}_2) - 2r(V(\bar{X}_1) * V(\bar{X}_2))}$$

where

$$V(\bar{X}_i) = (S.E.(\bar{X}_i))^2$$

and where r is the correlation coefficient between \bar{X}_1 and \bar{X}_2 . The correlation coefficient is generally no greater than 0.2 for CE estimates.

REFERENCES

1. U.S. Department of HEW, NCHS. Replication, as Approach to the Analysis of Data from Complex Surveys, by Philip J. McCarthy. Series 2, No. 14, Washington, D.C.: Gov't Printing Office 1966.
2. U.S. Department of HEW, NCHS. Pseudoreplication, Further Evaluation and Application of the Balanced Half-Sample Technique, by Philip J. McCarthy. Series 2, No. 31, Washington, D.C.: Gov't Printing Office, 1969.

VII. MICRODATA VERIFICATION AND ESTIMATION METHODOLOGY

This section is designed to help users become familiar with the microdata files. The following program gives users a benchmark to verify that their copy of the CD-ROM contains valid data, illustrate the methodology CE uses in producing publication tables, and offer an example of coding to access the data and produce a sample table. The program is written in SAS and utilizes the ASCII datasets available on this CD-ROM. A program written in SAS but utilizing the SAS datasets is also present on the CD-ROM but will not be referenced here. Refer to the table following the program to check output. (Note: CE data published by BLS may not match some values estimated using the microdata due to topcoding of data and CE publication programming methodology.) All variables and ranges referred to in the program are described in detail in Section III.E. DETAILED VARIABLE DESCRIPTIONS in this documentation.

This program produces a table of selected expenditures by income class of the Consumer Unit (CU). The first section of the program extracts the relevant variables from the FMLY files, while the second section extracts the expenditure and income data from the EXPN and DTAB files. These three datasets are then used along with the AGG and LABEL processing files to construct the sample table output. This output is the product of two SAS arrays. The values in one array are divided by the value in the other array to obtain weighted mean expenditures. The base, or denominator, for the division is a vector consisting of the weighted total population for the U.S. and selected income class categories. The numerator is a matrix of aggregate weighted costs for each line item in the table for the total U.S. population and each income class category.

It should be emphasized that this program has been written solely for the verification of the microdata and as an illustration of the CE estimation methodology. It should not be used for any other purpose.

Note: This program processes large amounts of data. If you are using a PC with limited capabilities it may be necessary to run this program in sections.

A. SAMPLE PROGRAM

```

1
2
3  %let y =96;
4
5  filename fmly1 "d:\diary&y\fmlyd&y.1.txt";
6  filename fmly2 "d:\diary&y\fmlyd&y.2.txt";
7  filename fmly3 "d:\diary&y\fmlyd&y.3.txt";
8  filename fmly4 "d:\diary&y\fmlyd&y.4.txt";
9
10 filename dtab1 "d:\diary&y\dtabd&y.1.txt";
11 filename dtab2 "d:\diary&y\dtabd&y.2.txt";
12 filename dtab3 "d:\diary&y\dtabd&y.3.txt";
13 filename dtab4 "d:\diary&y\dtabd&y.4.txt";
14
15 filename expn1 "d:\diary&y\expnd&y.1.txt";
16 filename expn2 "d:\diary&y\expnd&y.2.txt";
17 filename expn3 "d:\diary&y\expnd&y.3.txt";
18 filename expn4 "d:\diary&y\expnd&y.4.txt";
19
20 filename agg "d:\diary&y\aggd&y..txt";
21 filename labls "d:\diary&y\labeld&y..txt";
22
23
24
25 options linesize=153 pagesize=52 missing="";
26
27
28
29 data fmly1;
30   infile fmly1 lrecl=1549;
31   input @1 newid 8. @148 finlwt21 11.3
32     @1516 inclass $2.;

```

NOTE: The infile FMLY1 is:
 FILENAME=d:\diary96\fmlyd961.txt,
 RECFM=V,LRECL=1549

NOTE: 2135 records were read from the infile FMLY1.
 The minimum record length was 1549.
 The maximum record length was 1549.

NOTE: The data set WORK.FMLY1 has 2135 observations
 and 3 variables.

NOTE: The DATA statement used 41.68 seconds.

```

33  proc sort; by newid;
34

```

NOTE: The data set WORK.FMLY1 has 2135 observations
 and 3 variables.

NOTE: The PROCEDURE SORT used 0.82 seconds.

```

35  data fmly2;

```

Line 3 sets the year as a macro variable
 that can be used throughout the program.
 Lines 5-18 designate the location of the
 data on the cd-rom.

Lines 20-21 designate the location of the
 two auxiliary files.

Line 25 forces the output to be printed
 landscape.

Lines 29-51 pull in the necessary
 variables from the fmly files. Newid is the
 code given to a consumer unit each time
 they participate. Finlwt21 will be used to
 weight each consumer unit such that it
 represents some portion of the
 population. Inclass is a code that
 represents the range within which the
 consumer unit's annual income falls.

```

36  infile fmly2 lrecl=1549;

```

```
37   input @1 newid 8. @148 finlwt21 11.3
38       @1516 inclass $2.;
```

NOTE: The infile FMLY2 is:
FILENAME=d:\diary96\fmlyd962.txt,
RECFM=V,LRECL=1549

NOTE: 2481 records were read from the infile FMLY2.

The minimum record length was 1549.

The maximum record length was 1549.

NOTE: The data set WORK.FMLY2 has 2481 observations
and 3 variables.

NOTE: The DATA statement used 49.7 seconds.

```
39   proc sort; by newid;
40
```

NOTE: The data set WORK.FMLY2 has 2481 observations
and 3 variables.

NOTE: The PROCEDURE SORT used 0.6 seconds.

```
41   data fmly3;
42       infile fmly3 lrecl=1549;
43       input @1 newid 8. @148 finlwt21 11.3
44           @1516 inclass $2.;
```

NOTE: The infile FMLY3 is:
FILENAME=d:\diary96\fmlyd963.txt,
RECFM=V,LRECL=1549

NOTE: 2592 records were read from the infile FMLY3.

The minimum record length was 1549.

The maximum record length was 1549.

NOTE: The data set WORK.FMLY3 has 2592 observations
and 3 variables.

NOTE: The DATA statement used 53.92 seconds.

```
45   proc sort; by newid;
46
```

NOTE: The data set WORK.FMLY3 has 2592 observations
and 3 variables.

NOTE: The PROCEDURE SORT used 0.6 seconds.

```
47   data fmly4;
48       infile fmly4 lrecl=1549;
49       input @1 newid 8. @148 finlwt21 11.3
50           @1516 inclass $2.;
```

NOTE: The infile FMLY4 is:
FILENAME=d:\diary96\fmlyd964.txt,

RECFM=V,LRECL=1549

NOTE: 3568 records were read from the infile FMLY4.

The minimum record length was 1549.

The maximum record length was 1549.

NOTE: The data set WORK.FMLY4 has 3568 observations and 3 variables.

NOTE: The DATA statement used 1 minute 8.82 seconds.

```
51  proc sort; by newid;
52
53
```

NOTE: The data set WORK.FMLY4 has 3568 observations and 3 variables.

NOTE: The PROCEDURE SORT used 0.7 seconds.

```
54  data fmyall;
55  set fmy1 fmy2 fmy3 fmy4 ;
56  by newid;
57
58  uspop = finlwt21 / 4;
```

Lines 54-56 bring each of the 4 quarters of fmy datasets together.

Line 58 divides finlwt21 by 4 so that summing uspop later will yield the total U.S. population. (Since summing finlwt21 for each quarter will yield one U.S. population, this adjustment is necessary).

NOTE: The data set WORK.FMLYALL has 10776 observations and 4 variables.

NOTE: The DATA statement used 4.33 seconds.

```
59  proc sort; by newid;
60
```

NOTE: The data set WORK.FMLYALL has 10776 observations and 4 variables.

NOTE: The PROCEDURE SORT used 2.08 seconds.

```
61  proc summary nway data = fmyall (drop=finlwt21);
62  class inclass;
63  var uspop;
64  output out = newpop sum = popus;
```

Lines 61-73 create the total population weights by income group that will be used as the denominator in calculating the average annual expenditures later in the program and prints them.

NOTE: The data set WORK.NEWPOP has 10 observations and 4 variables.

NOTE: The PROCEDURE SUMMARY used 2.02 seconds.

```
65  proc transpose data = newpop out = transpop prefix
= pop;
66  var popus;
67
```

Lines 65-66 transpose the newpop dataset to match the format of the PUBRAY data set that it will be matched with later in the program.

NOTE: The data set WORK.TRANSPOP has 1 observations and 11 variables.

NOTE: The PROCEDURE TRANSPOSE used 0.55 seconds.

```
68  data subagg (drop = _name_);
69  set transpop;
```



```
70    popt = sum (of pop1-pop10);
71    popc = sum (of pop1-pop9);
```

NOTE: The data set WORK.SUBAGG has 1 observations and 12 variables.

NOTE: The DATA statement used 0.44 seconds.

Lines 68-71 take the transposed dataset and calculate popt, the all consumer units population, and popc, the all complete income reporters population.

```
72    proc print data=subagg;
73        title "Population Counts for 19&y";
74
75
76
```

NOTE: The PROCEDURE PRINT used 0.93 seconds.

```
77    data dtab1;
78        infile dtab1 lrecl=28;
79        input @1 newid 8. @9 ucc $6. @15 amount 12. ;
```

NOTE: The infile DTAB1 is:

```
    FILENAME=d:\diary96\dtabd961.txt,
    RECFM=V,LRECL=28
```

NOTE: 33758 records were read from the infile DTAB1.

The minimum record length was 28.

The maximum record length was 28.

NOTE: The data set WORK.DTAB1 has 33758 observations and 3 variables.

NOTE: The DATA statement used 16.25 seconds.

Lines 77-95 pull in the dtab files. Newid is the consumer unit code. Ucc is a code that represents the type of income variable. Amount is the value that corresponds to the ucc code.

```
80    proc sort; by newid;
81
```

NOTE: The data set WORK.DTAB1 has 33758 observations and 3 variables.

NOTE: The PROCEDURE SORT used 5.21 seconds.

```
82    data dtab2;
83        infile dtab2 lrecl=28;
84        input @1 newid 8. @9 ucc $6. @15 amount 12. ;
```

NOTE: The infile DTAB2 is:

```
    FILENAME=d:\diary96\dtabd962.txt,
    RECFM=V,LRECL=28
```

NOTE: 39695 records were read from the infile DTAB2.

The minimum record length was 28.

The maximum record length was 28.

NOTE: The data set WORK.DTAB2 has 39695 observations and 3 variables.

NOTE: The DATA statement used 18.73 seconds.

```
85    proc sort; by newid;
86
```

NOTE: The data set WORK.DTAB2 has 39695 observations and 3 variables.

NOTE: The PROCEDURE SORT used 5.92 seconds.

```
87  data dtab3;
88    infile dtab3 lrecl=28;
89    input @1 newid 8. @9 ucc $6. @15 amount 12. ;
```

NOTE: The infile DTAB3 is:

FILENAME=d:\diary96\dtabd963.txt,
RECFM=V,LRECL=28

NOTE: 41109 records were read from the infile DTAB3.

The minimum record length was 28.

The maximum record length was 28.

NOTE: The data set WORK.DTAB3 has 41109 observations and 3 variables.

NOTE: The DATA statement used 18.78 seconds.

```
90    proc sort; by newid;
91
```

NOTE: The data set WORK.DTAB3 has 41109 observations and 3 variables.

NOTE: The PROCEDURE SORT used 6.58 seconds.

```
92  data dtab4;
93    infile dtab4 lrecl=28;
94    input @1 newid 8. @9 ucc $6. @15 amount 12. ;
```

NOTE: The infile DTAB4 is:

FILENAME=d:\diary96\dtabd964.txt,
RECFM=V,LRECL=28

NOTE: 56019 records were read from the infile DTAB4.

The minimum record length was 28.

The maximum record length was 28.

NOTE: The data set WORK.DTAB4 has 56019 observations and 3 variables.

NOTE: The DATA statement used 25.53 seconds.

```
95    proc sort; by newid;
96
```

NOTE: The data set WORK.DTAB4 has 56019 observations and 3 variables.

NOTE: The PROCEDURE SORT used 7.91 seconds.

```

97  data dtab(rename=(amount=cost));
98    set dtab1 dtab2 dtab3 dtab4 ;
99    by newid;

```

NOTE: The data set WORK.DTAB has 170581 observations and 3 variables.
 NOTE: The DATA statement used 26.85 seconds.

Lines 97-99 bring the 4 quarters of dtab datasets together. The variable amount is renamed cost so that it can be merged with the expn datasets later in the program.

```

100  proc sort; by newid;

```

NOTE: The data set WORK.DTAB has 170581 observations and 3 variables.
 NOTE: The PROCEDURE SORT used 28.17 seconds.

```

101  proc datasets;

```

-----Directory-----

-

Libref: WORK
 Engine: V612
 Physical Name:

C:\SAS612\SASWORK\#TD17925

Lines 101-102 delete from memory the datasets that are no longer necessary for processing.

Memtype	Indexes	#	Name
---			-----
		1	DTAB DATA
		2	DTAB1 DATA
		3	DTAB2 DATA
		4	DTAB3 DATA
		5	DTAB4 DATA
		6	FMLY1 DATA
		7	FMLY2 DATA
		8	FMLY3 DATA
		9	FMLY4 DATA
		10	FMLYALL
DATA		11	NEWPOP
DATA		12	SUBAGG
DATA		13	TRANSPop
DATA			
101	delete dtab1 dtab2 dtab3 dtab4 ;		
102			
103			

NOTE: Deleting WORK.DTAB1 (memtype=DATA).
 NOTE: Deleting WORK.DTAB2 (memtype=DATA).
 NOTE: Deleting WORK.DTAB3 (memtype=DATA).
 NOTE: Deleting WORK.DTAB4 (memtype=DATA).
 NOTE: The PROCEDURE DATASETS used 0.98 seconds.

```

104  data expn1;
105    infile expn1 lrecl=40;

```

106 input @1 newid 8. @35 ucc \$6. @10 cost 12.5 ;

NOTE: The infile EXPN1 is:

FILENAME=d:\diary96\expnd961.txt,
RECFM=V,LRECL=40

NOTE: 89058 records were read from the infile EXPN1.

The minimum record length was 40.

The maximum record length was 40.

NOTE: The data set WORK.EXP1 has 89058
observations and 3 variables.

NOTE: The DATA statement used 52.28 seconds.

Lines 104-122 pull in the expn files.
Newid is the consumer unit code. Ucc is
the code designating the type of
expenditure. Cost is the amount of the
expenditure.

107 proc sort; by newid;

108

NOTE: The data set WORK.EXP1 has 89058
observations and 3 variables.

NOTE: The PROCEDURE SORT used 16.85 seconds.

109 data expn2;

110 infile expn2 lrecl=40;

111 input @1 newid 8. @35 ucc \$6. @10 cost 12.5 ;

NOTE: The infile EXPN2 is:

FILENAME=d:\diary96\expnd962.txt,
RECFM=V,LRECL=40

NOTE: 107656 records were read from the infile EXPN2.

The minimum record length was 40.

The maximum record length was 40.

NOTE: The data set WORK.EXP2 has 107656
observations and 3 variables.

NOTE: The DATA statement used 1 minute 7.61 seconds.

112 proc sort; by newid;

113

NOTE: The data set WORK.EXP2 has 107656
observations and 3 variables.

NOTE: The PROCEDURE SORT used 20.98 seconds.

114 data expn3;

115 infile expn3 lrecl=40;

116 input @1 newid 8. @35 ucc \$6. @10 cost 12.5 ;

NOTE: The infile EXPN3 is:

FILENAME=d:\diary96\expnd963.txt,
RECFM=V,LRECL=40

NOTE: 111359 records were read from
the infile EXPN3.

The minimum record length was 40.
The maximum record length was 40.
NOTE: The data set WORK.EXP3 has 111359
observations and 3 variables.
NOTE: The DATA statement used 1 minute 9.65 seconds.

```
117  proc sort; by newid;  
118
```

NOTE: The data set WORK.EXP3 has 111359
observations and 3 variables.
NOTE: The PROCEDURE SORT used 22.01 seconds.

```
119  data expn4;  
120  infile expn4 lrecl=40;  
121  input @1 newid 8. @35 ucc $6. @10 cost 12.5 ;
```

NOTE: The infile EXPN4 is:
FILENAME=d:\diary96\expnd964.txt,
RECFM=V,LRECL=40

NOTE: 151625 records were read from the infile EXPN4.
The minimum record length was 40.
The maximum record length was 40.
NOTE: The data set WORK.EXP4 has 151625
observations and 3 variables.
NOTE: The DATA statement used 1 minute 39.57 seconds.

```
122  proc sort; by newid;  
123
```

NOTE: The data set WORK.EXP4 has 151625
observations and 3 variables.
NOTE: The PROCEDURE SORT used 27.19 seconds.

```
124  data expn;  
125  set expn1 expn2 expn3 expn4 ;  
126  by newid;  
127  if cost > 0;
```

Lines 124-128 bring all 4 quarters of the
expn files together.

NOTE: The data set WORK.EXP4 has 459698
observations and 3 variables.
NOTE: The DATA statement used 1 minute 23.37 seconds.

```
128  proc sort; by newid;
```

NOTE: The data set WORK.EXP4 has 459698
observations and 3 variables.
NOTE: The PROCEDURE SORT used 1 minute 24.35
seconds.

```
129  proc datasets;
```

-----Directory-----

Lines 129-130 delete from memory the datasets no longer needed for processing.

-
Libref: WORK
Engine: V612
Physical Name:
C:\SAS612\SASWORK\#TD17925

Memtype	Indexes	#	Name
---			-----
		1	DTAB DATA
		2	EXPN DATA
		3	EXPN1 DATA
		4	EXPN2 DATA
		5	EXPN3 DATA
		6	EXPN4 DATA
		7	FMLY1 DATA
		8	FMLY2 DATA
		9	FMLY3 DATA
		10	FMLY4 DATA
		11	FMLYALL DATA
		12	NEWPOP DATA
		13	SUBAGG DATA
		14	TRANSPOP DATA

129 delete expn1 expn2 expn3 expn4;
130
131
132

NOTE: Deleting WORK.EXP1 (memtype=DATA).
NOTE: Deleting WORK.EXP2 (memtype=DATA).
NOTE: Deleting WORK.EXP3 (memtype=DATA).
NOTE: Deleting WORK.EXP4 (memtype=DATA).
NOTE: The PROCEDURE DATASETS used 0.77 seconds.

133 data expend ;
134 set dtab expn;
135 by newid;

NOTE: The data set WORK.EXPEND has 630279 observations and 3 variables.

NOTE: The DATA statement used 1 minute 48.64 seconds.

136 proc sort; by newid;

NOTE: The data set WORK.EXPEND has 630279 observations and 3 variables.

Lines 133-136 pull the dtab and expn files together.

NOTE: The PROCEDURE SORT used 2 minutes 35.6 seconds.

137 proc datasets;

-----Directory-----

Libref: WORK
Engine: V612
Physical Name:

Lines 137-138 delete from memory the datasets no longer needed for processing.

C:\SAS612\SASWORK\#TD17925

Memtype	Indexes	#	Name
---			-----
		1	DTAB DATA
		2	EXPEND
DATA			
		3	EXPN DATA
		4	FMLY1 DATA
		5	FMLY2 DATA
		6	FMLY3 DATA
		7	FMLY4 DATA
		8	FMLYALL
DATA			
		9	NEWPOP
DATA			
		10	SUBAGG
DATA			
		11	TRANSPop
DATA			
137	delete dtab expn;		
138			

NOTE: Deleting WORK.DTAB (memtype=DATA).
NOTE: Deleting WORK.EXPND (memtype=DATA).
NOTE: The PROCEDURE DATASETS used 0.38 seconds.

139 data pubfile (drop= uspop) ;
140 merge fmlyall (in = infam)
141 expnd (in = inexp)
142 ;
143 by newid ;
144 if not inexp then delete;
145 if cost='.' then cost=0;
146
147 wcost = finlwt21 * cost/4;
148

Lines 139-147 merge the fmlyall and expnd datasets together and check the cost variable to make sure that there are no missing values.

NOTE: Character values have been converted to numeric values at the places given by: (Line):(Column).145:13
NOTE: The data set WORK.PUBFILE has 630279 observations and 6 variables.
NOTE: The DATA statement used 3 minutes 21.35 seconds.

Line 147 weights the cost variable up to the population level that the consumer unit represents. (Finlwt21 will sum to one U.S. population each quarter so it must be adjusted by dividing by 4 so that summing all 4 quarters will yield the U.S. population).

149 proc summary nway data =
pubfile (drop=newid);

```

150 class ucc inclass;
151 var wcost ;
152 output out = aggcst sum = ;
153

```

Lines 149-152 sum the weighted costs for the consumer units for each ucc by income group and outputs this as a new dataset called aggcst.

NOTE: The data set WORK.AGGCST has 4688 observations and 5 variables.

NOTE: The PROCEDURE SUMMARY used 4 minutes 2.16 seconds.

```

154 proc datasets;

```

-----Directory----

```

Libref:   WORK
Engine:   V612
Physical Name:

```

Lines 154-155 delete from memory any datasets that are no longer needed for processing.

```

C:\SAS612\SASWORK\#TD17925

```

Name

Memtype Indexes

```

1 AGGCST  DATA
2 EXPEND  DATA
3 FMLY1   DATA
4 FMLY2   DATA
5 FMLY3   DATA
6 FMLY4   DATA
7 FMLYALL DATA
8 NEWPOP  DATA
9 PUBFILE DATA
10 SUBAGG DATA
11 TRANSPOP DATA

```

```

155 delete expend pubfile;
156

```

NOTE: Deleting WORK.EXPEND (memtype=DATA).

NOTE: Deleting WORK.PUBFILE (memtype=DATA).

NOTE: The PROCEDURE DATASETS used 0.44 seconds.

```

157 data aggray1 (drop = inclass _type_ _freq_ wcost);
158 set aggcst;
159 by ucc ;
160 array trncost grp1-grp10;
161 retain grp1-grp10;
162 if first.ucc then do over trncost;
163     trncost = 0;
164 end;
165 _I_=inclass;
166 trncost=wcost;
167 if last.ucc then output;

```

Lines 157-167 create the variables grp1-grp10 that will designate the income groups and then places the weighted cost, or expenditure, data into the appropriate new variable.

NOTE: Character values have been converted to numeric values at the places given by: (Line):(Column).

165:13

NOTE: The data set WORK.AGGRAY1 has 529 observations and 11 variables.

NOTE: The DATA statement used 2.95 seconds.

```
169 data agfile;
170     infile agg missover pad;
171     input @3 ucc $6.
172         @15 line $6.;
```

Lines 169-174 pull in the file that dictates how each ucc will be summed for aggregation.

NOTE: The infile AGG is:

```
FILENAME=c:\tmp\aggd96.txt,
RECFM=V,LRECL=256
```

NOTE: 595 records were read from the infile AGG.

The minimum record length was 0.

The maximum record length was 26.

NOTE: The data set WORK.AGFILE has 595 observations and 2 variables.

NOTE: The DATA statement used 0.66 seconds.

```
173 proc sort data = agfile;
174     by ucc ;
175
```

NOTE: The data set WORK.AGFILE has 595 observations and 2 variables.

NOTE: The PROCEDURE SORT used 0.39 seconds.

```
176 data pubray ;
177     merge aggray1 (in = inray)
178         agfile (in = inagg);
179     by ucc;
180     if inray and inagg;
181
```

Lines 176-180 merge the dataset containing the weighted costs and the agfile. The agfile will give all costs a code called line that will be used for aggregation.

NOTE: The data set WORK.PUBRAY has 583 observations and 12 variables.

NOTE: The DATA statement used 0.7 seconds.

```
182 proc summary nway data = pubray;
183     class line;
184     var grp1-grp10;
185     output out =aggsum sum = ;
186
```

Lines 182-185 sum the weighted costs for each income group (grp1-grp10) by line and output this into a new dataset called aggsum.

NOTE: The data set WORK.AGGSUM has 67 observations and 13 variables.

NOTE: The PROCEDURE SUMMARY used 0.59 seconds.

```
187 data cstpop1 (drop = _type_ _freq_ popt popc pop1-  
pop10);  
188   if _n_ = 1 then set subagg;  
189   set aggsum;  
190   grpt = sum (of grp1-grp10);  
191   grpc = sum (of grp1-grp9);  
192   array ex grpt grpc grp1-grp10;  
193   array wt popt popc pop1-pop10;  
194   do over ex;  
195     ex = ex/wt;  
196   end;  
197
```

Lines 187-196 create two arrays. One array is a vector from the subagg dataset that contains the population counts (popt, popc pop1-pop10). The other is a matrix of the weighted costs by income group. The costs are divided by the population counts.

NOTE: The data set WORK.CSTPOP1 has 67 observations and 13 variables.

NOTE: The DATA statement used 0.88 seconds.

```
198 data numcus (rename=(popt=grpt popc=grpc  
pop1=grp1 pop2=grp2  
199           pop3=grp3 pop4=grp4 pop5=grp5  
pop6=grp6  
200           pop7=grp7 pop8=grp8 pop9=grp9  
pop10=grp10));  
201   set subagg;  
202   line = '000000';  
203
```

Lines 198-206 give the population counts a line value so that they can be printed as part of the final output, and then brings them together with the summed cost dataset that was calculated with the arrays.

NOTE: The data set WORK.NUMCUS has 1 observations and 13 variables.

NOTE: The DATA statement used 0.22 seconds.

```
204 data cstpop;  
205   set numcus cstpop1;  
206   by line;  
207
```

NOTE: The data set WORK.CSTPOP has 68 observations and 13 variables.

NOTE: The DATA statement used 0.28 seconds.

```
208 data addlab ;  
209   infile labls missover pad;  
210   input @1 line $6. @10 title $char40.;
```

NOTE: The infile LABLS is:

```
FILENAME=c:\tmp\labld96.txt,  
RECFM=V,LRECL=256
```

NOTE: 67 records were read from the infile LABLS.

The minimum record length was 56.

The maximum record length was 58.

Lines 208-211 pull in the label file that will put titles on the final output.

NOTE: The data set WORK.ADDLAB has 67 observations and 2 variables.

NOTE: The DATA statement used 0.44 seconds.

```
211 proc sort; by line;
212
```

NOTE: The data set WORK.ADDLAB has 67 observations and 2 variables.

NOTE: The PROCEDURE SORT used 0.33 seconds.

```
213 data pubtab (drop = line);
214 merge cstpop (in = inline)
215 addlab (in = inlabl);
216 by line;
217 if not inlabl then delete;
218
```

Lines 213-238 merge the summed cost dataset with the titles for printing. The output is formatted and the income groups are given labels. Note that not all groups are printed – the incomplete reporters (grp10) and all consumer units (grpt).

NOTE: The data set WORK.PUBTAB has 67 observations and 13 variables.

NOTE: The DATA statement used 0.33 seconds.

```
219 proc print split='*' uniform;
220 label
221 grpt=' All* Consumer* Units*'
222 grpc=' Total* Complete*Reporting*'
223 grp1=' Less* Than* $5,000*'
224 grp2=' $5,000* To* $9,999*'
225 grp3=' $10,000* To* $14,999*'
226 grp4=' $15,000* To* $19,999*'
227 grp5=' $20,000* To* $29,999*'
228 grp6=' $30,000* To* $39,999*'
229 grp7=' $40,000* To* $49,999*'
230 grp8=' $50,000* To* $69,999*'
231 grp9=' $70,000* And* Over*'
232 grp10='Incomplete*
Income*Reporters*';
233 format title $char40.;
234 format grpt grpc grp1-grp10 comma9.2;
235 id title;
236 var grpc grp1-grp9;
237 title "CE Microdata Diary Survey Average Weekly
Expenditures, for Calendar Year 19&y by Income";
238 title2 ' ';
239
240 run;
```

NOTE: At least one W.D format was too small for the number to be printed. The decimal may be shifted by the "BEST" format.

NOTE: The PROCEDURE PRINT used 1.26 seconds.

B. OUTPUT

The following observation shows the contents of the SUBAGG data set created in lines 68-73. It represents the weighted number of CUs in each INCLASS category as well as for the total population and the population of complete income reporters.

Population Counts for 1996												1
OBS	POP1	POP2	POP3	POP4	POP5	POP6	POP7	POP8	POP9	POP10	POPT	POPC
1	3728449.74	7942865.69	7977661.84	7225006.21	12716869.86	10298286.38	7410972.32	10905541.03	12028784.15	23740223.88	103974661.12	80234437.23

The table that follows represents printed output of the PUBTAB dataset.

CE Microdata Diary Survey Average Weekly Expenditures, for Calendar Year 1996 by Income											2
TITLE	Total Complete Reporting	Less Than \$5,000	\$5,000 To \$9,999	\$10,000 To \$14,999	\$15,000 To \$19,999	\$20,000 To \$29,999	\$30,000 To \$39,999	\$40,000 To \$49,999	\$50,000 To \$69,999	\$70,000 And Over	
Number of consumer units	80234437	3728449.7	7942865.7	7977661.8	7225006.2	12716870	10298286	7410972.3	10905541	12028784	
Income before taxes	40,480.73	1,871.76	7,569.74	12,261.63	17,315.63	24,512.42	34,473.70	44,494.83	58,737.94	109808.26	
Income after taxes.....	36,996.38	1,887.98	7,434.38	12,012.86	16,815.05	23,328.78	31,734.17	40,798.51	53,232.10	97,982.67	
Age of reference person.....	47.36	39.24	58.21	52.76	54.05	47.08	43.09	42.60	42.71	46.20	
Average number in consumer unit:											
Persons	2.54	1.74	1.70	2.08	2.12	2.50	2.69	2.83	3.05	3.16	
Children under 18.....	0.71	0.43	0.42	0.55	0.49	0.70	0.80	0.89	0.92	0.86	
Persons 65 and over.....	0.30	0.17	0.56	0.52	0.60	0.41	0.21	0.13	0.11	0.11	
Earners.....	1.43	0.94	0.51	0.83	0.96	1.34	1.54	1.77	1.99	2.18	
Vehicles.....	1.69	0.98	0.80	1.11	1.44	1.63	1.78	1.97	2.13	2.46	
Percent distribution:											
Male	58.34	43.98	29.51	43.78	48.23	59.70	64.84	65.91	71.76	73.76	
Female	41.66	56.02	70.49	56.22	51.77	40.30	35.16	34.09	28.24	26.24	
Homeowner	61.85	34.92	41.30	47.00	58.82	58.73	56.90	71.25	74.55	85.64	
Renter	36.33	64.07	57.65	50.84	39.73	39.71	40.47	26.81	24.09	11.83	
Black	11.88	15.25	19.99	15.98	12.51	13.01	12.07	9.31	8.88	5.34	
White and other	88.12	84.75	80.01	84.02	87.49	86.99	87.93	90.69	91.12	94.66	
Elementary education	8.43	12.07	24.95	16.28	12.17	7.53	6.15	2.37	2.56	0.92	
High school education	38.10	39.98	44.97	49.94	51.42	44.65	39.90	36.06	29.74	17.47	
College education	52.94	47.95	29.10	33.48	34.88	47.22	53.49	61.22	67.55	81.18	
Never attended and other	0.54	0.00	0.98	0.30	1.52	0.60	0.47	0.35	0.15	0.43	
CE Microdata Diary Survey Average Weekly Expenditures, for Calendar Year 1996 by Income											3
Total	Less	\$5,000	\$10,000	\$15,000	\$20,000	\$30,000	\$40,000	\$50,000	\$70,000		

TITLE	Complete Reporting	Than \$5,000	To \$9,999	To \$14,999	To \$19,999	To \$29,999	To \$39,999	To \$49,999	To \$69,999	And Over
At least one vehicle owned	88.98	61.64	60.83	76.73	89.47	93.43	97.79	97.57	97.49	98.65
Food, total.....	86.42	43.94	45.86	60.27	64.60	74.02	87.49	97.42	111.33	139.64
Food at home.....	56.67	31.99	36.82	45.09	46.03	51.33	56.21	62.80	68.78	82.76
Cereals and cereal products.....	3.31	2.13	2.18	2.87	2.52	3.05	3.29	3.62	3.86	4.76
Bakery products.....	5.68	3.17	3.41	3.94	4.62	4.85	5.50	6.48	7.28	8.85
Beef.....	4.22	2.77	2.83	3.86	3.55	3.93	3.93	4.71	5.45	5.39
Pork.....	3.08	1.76	2.32	2.77	2.77	3.04	3.18	3.30	3.62	3.72
Other meats.....	1.95	1.20	1.30	1.69	1.45	1.78	1.81	2.00	2.46	2.88
Poultry.....	2.90	1.89	1.90	2.49	2.42	2.53	2.87	3.24	3.39	4.20
Fish and seafood.....	1.76	1.07	1.10	1.44	1.30	1.56	1.50	1.72	1.90	3.22
Eggs.....	0.67	0.42	0.58	0.69	0.65	0.72	0.73	0.63	0.67	0.74
Fresh milk and cream.....	2.67	1.58	1.83	2.24	2.17	2.68	2.90	3.11	3.20	3.21
Other dairy products.....	3.67	1.85	2.06	2.49	2.82	3.09	3.81	4.01	4.63	6.01
Fresh fruits.....	3.04	1.34	1.71	2.52	2.75	2.72	3.04	3.08	3.67	4.71
Fresh vegetables.....	2.84	1.34	1.94	2.49	2.54	2.53	2.75	2.75	3.12	4.51
Processed fruits.....	2.21	1.10	1.48	1.65	1.84	2.07	2.09	2.60	2.57	3.29
Processed vegetables.....	1.60	0.97	1.15	1.28	1.42	1.55	1.57	1.63	1.93	2.20
Sugar and other sweets.....	2.33	1.21	1.55	1.74	1.88	2.16	2.39	2.65	2.96	3.24
Fats and oils.....	1.67	1.07	1.36	1.58	1.47	1.61	1.65	1.69	1.96	2.01
Miscellaneous foods.....	7.99	4.30	4.67	5.54	6.05	6.88	7.80	9.60	10.23	12.41
Nonalcoholic beverages.....	5.07	2.82	3.45	3.81	3.81	4.58	5.39	5.96	5.87	7.41
Food away from home.....	29.75	11.95	9.04	15.18	18.57	22.69	31.29	34.61	42.56	56.88
Alcoholic beverages.....	5.73	3.69	2.16	2.90	3.36	4.07	6.64	7.02	6.08	11.92
Housing	60.19	29.32	34.33	39.70	46.68	55.77	58.62	65.12	74.79	98.28
Fuel and utilities	41.40	21.50	28.23	29.64	33.27	40.31	43.83	44.76	49.06	58.98
Household operations	0.24	0.09	0.09	0.04	0.09	0.08	0.19	0.25	0.29	0.79
Housekeeping supplies	9.77	4.95	4.46	5.84	8.47	8.15	8.14	11.43	12.23	17.99
Household furnishings and equipment ...	8.78	2.78	1.55	4.17	4.85	7.22	6.46	8.68	13.22	20.52
Apparel and services.....	31.47	15.13	12.30	14.71	20.76	24.72	25.64	31.16	44.48	67.28
Men, 16 and over.....	5.41	2.33	1.30	2.60	3.31	4.48	4.85	6.25	7.87	10.91
Boys, 2 to 15.....	0.85	0.35	0.44	0.37	0.25	0.64	0.68	0.99	1.36	1.79
Women, 16 and over.....	11.92	7.13	5.76	4.88	7.73	9.25	8.78	10.61	17.52	25.91
Girls, 2 to 15.....	1.21	0.27	0.22	0.43	0.67	0.90	0.99	1.41	2.53	2.16
Children under 2.....	1.49	0.38	0.76	0.76	0.87	1.42	0.98	2.80	2.05	2.38
Footwear.....	5.90	2.66	2.89	3.49	4.26	5.29	5.30	6.09	8.34	10.29
Other apparel products and services...	4.69	1.99	0.93	2.19	3.67	2.73	4.05	3.01	4.82	13.83
Transportation	20.88	8.49	8.19	12.05	15.31	20.00	23.14	24.36	29.01	31.77
Non-prescription drugs and supplies ...	2.54	1.24	1.61	2.17	2.51	2.50	2.50	2.69	3.05	3.36
Entertainment	9.19	3.96	2.28	4.11	3.87	6.50	9.42	10.77	13.07	20.10
Radios, sound equipmnet	0.53	0.00	0.23	0.04	0.60	0.53	0.79	0.30	1.22	0.45
Pet food and supplies	3.08	1.22	1.19	2.04	1.29	2.11	3.20	2.57	4.25	6.83
Toys, games, playground equipment	2.59	1.41	0.58	1.14	1.23	1.88	2.85	2.85	3.77	5.34
Other entertainment supplies, equipment	3.00	1.33	0.28	0.89	0.76	1.98	2.59	5.04	3.83	7.47
Personal care products and services.....	4.84	1.99	1.79	3.08	2.92	4.09	5.13	6.03	6.16	8.66
Miscellaneous	2.77	0.91	1.60	1.22	2.76	2.49	3.05	2.70	2.77	5.27

VIII. DESCRIPTION OF THE SURVEY

The CE program consists of two separate components, each with its own questionnaire and independent sample:

1) A Diary or recordkeeping survey completed by sample consumer units for two consecutive 1-week periods; the sample is spread across a 12-month period.

2) An Interview panel survey in which each CU in the sample is interviewed every 3 months over five consecutive quarters. New panels are initiated every month of the year.

Data are collected by the Bureau of the Census under contract with BLS. All data collected in both surveys are subject to Bureau of the Census confidentiality requirements which prevent the disclosure of the CU member's identity.

The Diary Survey collects expenditure data for items purchased each day over two one-week periods. This survey is designed to collect expenditure data for small, frequently purchased items such as food, beverages, food consumed away from home, gasoline, housekeeping supplies, nonprescription drugs and medical supplies, and personal care products and services.

A Household Characteristics Questionnaire is completed to record demographic and family characteristics data pertaining to age, sex, race, marital status, CU relationships, work experience, and income of each CU member. The expenditure collection instrument is a self-reporting, product-oriented diary on which respondents record all expenses for two consecutive one-week periods. It is divided by day of purchase and by broad classification of goods and services, a breakdown designed to aid the respondents when recording daily purchases.

At the beginning of the two-week collection period, the interviewer uses the Household Characteristics Questionnaire to record information pertaining to household members. Also at this time, a diary for the first week is left with the participating CU. At the completion of the first week, the interviewer picks up the diary, reviews the entries, clarifies any questions, and leaves a second diary for the following week. At the end of the second week, the diary is picked up and reviewed. At this point, the interviewer again uses the Household Characteristics Questionnaire to collect information on employment and earnings of the CU members. These data, along with the other household characteristics information, permits data users to classify sample units for research purposes, and allows BLS to adjust nonresponses by CUs who do not cooperate in the survey.

IX. DATA COLLECTION AND PROCESSING

In addition to its data collection duties, the Census Bureau is responsible for field editing and coding, consistency checking, quality control, and data transmittal to BLS. BLS performs additional review and editing procedures in preparing the data for publication and release.

A. BUREAU OF THE CENSUS ACTIVITIES

Data collection activities have been conducted by the Census Bureau on a continuing basis since October 1979. Due to differences in format and design, the Diary Survey and the Interview Survey data are collected and processed separately.

Preliminary Diary survey processing carried out by the Bureau of the Census includes clerical data edits and adjustments. Upon completion by respondents, the diaries are returned to the regional offices, where they undergo a cursory field edit. They are then sent to the Data Preparation Division in Jeffersonville, Indiana, where more detailed edits are done for completeness and consistency. Codes are also assigned to reported expenditure items and to demographic characteristics of CU members.

After clerical processing, the data are keyed and transmitted to the Census Processing Center in Washington, D.C. Here computer programs are run which derive CU weights based on BLS specifications, impute demographic and work experience characteristics when missing or invalid, and apply appropriate sales taxes to expenditure items. Final tapes of the edited and coded data are then transmitted to the BLS on a monthly basis.

B. BUREAU OF LABOR STATISTICS ACTIVITIES

Upon receipt of the data from the Bureau of the Census, BLS conducts an extensive review to ensure that severe data aberrations are corrected. The review takes place in several stages: a review of counts, weighted means, and unweighted means by region; a review of family relationship coding inconsistencies; a review of selected extreme values for expenditure and income categories; and a verification of the various data transformations.

Cases of extreme data values are investigated by viewing diaries on microfilm. Errors discovered through this procedure are corrected prior to release of the data for public use.

Two major types of data adjustment routines--imputation and allocation--are carried out to improve and classify the estimates derived from the Diary Survey. Data imputation routines correct for missing or invalid entries among selected CU characteristics fields. No imputations are performed for income fields. Allocation routines are applied when respondents provided insufficient expenditure detail to meet tabulation requirements. For example, reports of combined expenditures for fuels and utilities are allocated among gas, electricity, and other items in this group. To analyze the effects of these adjustments, tabulations are made before and after the data adjustments.

X. SAMPLING STATEMENT

A. SURVEY SAMPLE DESIGN

Samples for the Consumer Expenditure Survey are national probability samples of households designed to be representative of the total U. S. civilian population. Eligible population includes all civilian noninstitutional persons.

The first step in sampling is the selection of primary sampling units (PSU's), which consists of counties (or parts thereof) or groups of counties. The set of sample PSU's used for the 1996 and 1997 samples is composed of 105 areas. This new design classifies the PSU's into four categories:

- 31 "A" certainty PSU's are Metropolitan Statistical Areas (MSA's) with a population greater than 1.5 million.
- 46 "B" PSU's, are medium-sized MSA's.
- 10 "C" PSU's are nonmetropolitan areas which are included in the CPI.
- 18 "D" PSU's are nonmetropolitan areas where only the urban population data will be included in the CPI.

The sampling frame (that is, the list from which housing units were chosen) for the 1996 and 1997 surveys were generated from the 1990 Population Census 100-percent-detail file. The sampling frame was augmented by new construction permits and by techniques used to eliminate recognized deficiencies in census coverage. All Enumeration Districts (ED's) from the Census that fail to meet the criterion for good addresses for new construction, and all ED's in nonpermit-issuing areas are grouped into the area segment frame.

To the extent possible, an unclustered sample of units is selected within each PSU. This lack of clustering is desirable because the sample size of the Diary Survey is small relative to other surveys, while the intraclass correlations for expenditure characteristics are relatively large. This suggests that any clustering of the sample units could result in an unacceptable increase in the within-PSU variance and, as a result, the total variance.

Each selected sample unit is requested to keep two 1-week diaries of expenditures over consecutive weeks. The earliest possible day for placing a diary with a household is predesignated with each day of the week having an equal chance to be the first of the reference week. The diaries are evenly spaced throughout the year. During the last 6 weeks of the year, however, the Diary Survey sample is supplemented to twice its normal size to increase the reporting of types of expenditures unique to the holidays.

B. COOPERATION LEVELS

The annual target sample size at the United States level for the Diary Survey is 6,050 participating sample units. To achieve this target the total estimated work load is 8,180 sample units. The difference between participating sample units and total work load is due to refusals, vacancies, or nonexistence of sample unit address.

Each participating sample unit selected is asked to keep two 1-week diaries. Each diary is treated independently, so the number of sample units used in the following calculations is actually twice the number of housing units sampled.

The response rate for the 1996 Diary Survey is 73.8% as shown below. This response rate refers to all diaries in the year.

Number of diaries designated for the survey	Type B or C ineligible cases	<i>Eligible housing unit interviews</i>		
		Number of potential diaries	Type A nonresponse	Total respondent interviews
18,219	3,613	14,606	3,830	10,776

Type B or C cases are housing units that are vacant, nonexistent, or ineligible for Diary. Type A nonresponses are housing units which the Interviewers were unable to contact or the respondents refused to participate in the survey. The response rate stated above is based only on the eligible housing units (i.e., the designated sample cases less type B and type C ineligible cases).

C. WEIGHTING

Each CU included in the CE represents a given number of CUs in the U.S. population, which is considered to be the universe. The translation of sample families into the universe of families is known as weighting. However, since the unit of analysis for the CE is a consumer unit, the weighting was performed at the CU level. Several factors are involved in determining the weight for each consumer unit for which a diary is obtained.

There are four basic steps in the weighting procedure:

- 1) The basic weight is assigned to an address and is the inverse of the probability of selection of the housing unit.
- 2) A weight control factor is applied to each diary if subsampling is performed in the field.
- 3) A noninterview adjustment is made for units where data could not be collected from occupied housing units. The adjustment is performed as a function of region, housing tenure, family size and race.
- 4) A final adjustment is performed to adjust the sample estimates to national population controls derived from the Current Population Survey. The adjustments are made based on the CU composition and based on the CU as a whole. The weight for the CU is adjusted for individuals within the CU to meet the controls for the 14 age/race categories, 4 regions, and 4 region/urban categories. The CU weight is also adjusted to meet the control for total number of CUs and total number of Owner CUs. The weighting procedure uses an iterative process to ensure that the sample estimates will meet all the population controls.

NOTE: The weight for a consumer unit (CU) can be different for each week in which the CU participates in the survey because a diary respondent CU may represent a different number of CUs with similar characteristics each week.

D. STATE IDENTIFIER

Since the CE is not designed to produce state-level estimates, summing the consumer unit weights by state will not yield state population totals. To understand why, a brief explanation of how CE weights are constructed follows.

A consumer unit's basic weight reflects its probability of selection among a group of primary sampling units of similar characteristics. For example, sample units in an urban nonmetropolitan area in California may represent similar areas in Wyoming and Nevada. Among other adjustments, consumer units are post-stratified nationally by sex-age-race. For example, the weights of consumer units containing a black male, age 16-24 in Alabama, Colorado, or New York, are all adjusted equivalently. Therefore, weighted population state totals will not match population totals calculated from other surveys that are designed to represent state data.

To summarize, the CE sample was not designed to produce precise estimates for individual states. Although state-level estimates that are unbiased in a repeated sampling sense can be calculated for various statistical measures, such as means and aggregates, their estimates will generally be subject

to large variances. Additionally, a particular state-population estimate from the CE sample may be far from the true state-population estimate.

XI. INTERPRETING THE DATA

Several factors should be considered when interpreting the expenditure data. The average expenditure for an item may be considerably lower than the expenditure by those consumer units that purchased the item. The less frequently an item is purchased, the greater the difference between the average for all consumer units and the average of those purchasing. (See Section V.B. for ESTIMATION OF TOTAL AND MEAN EXPENDITURES). Also, an individual consumer unit may spend more or less than the average, depending on its particular characteristics. Factors such as income, age of family members, geographic location, taste and personal preference also influence expenditures. Furthermore, even within groups with similar characteristics, the distribution of expenditures varies substantially.

Expenditures reported are the direct out-of-pocket expenditures. Indirect expenditures, which may be significant, may be reflected elsewhere. For example, rental contracts often include utilities. Renters with such contracts would record no direct expense for utilities, and therefore, appear to have no utility expenses. Employers or insurance companies frequently pay other costs. Consumer units with members whose employers pay for all or part of their health insurance or life insurance would have lower direct expenses for these items than those who pay the entire amount themselves.

These points should be considered when relating reported averages to individual circumstances.

XII. APPENDIX 1--GLOSSARY

Population

The civilian noninstitutional population of the United States as well as that portion of the institutional population living in the following group quarters: Boarding houses, housing facilities for students and workers, staff units in hospitals and homes for the aged, infirm, or needy, permanent living quarters in hotels and motels, and mobile home parks. Urban population is defined as all persons living in a Metropolitan Statistical Area (MSA) and in urbanized areas and urban places of 2,500 or more persons outside of MSA's. Urban, defined in this survey, includes the rural populations within an MSA. The general concept of an MSA is one of a large population nucleus together with adjacent communities which have a high degree of economic and social integration with that nucleus. Rural population is defined as all persons living outside of an MSA and within an area with less than 2,500 persons.

Consumer unit (CU)

A consumer unit comprises either: (1) all members of a particular household who are related by blood, marriage, adoption, or other legal arrangements; (2) a person living alone or sharing a household with others or living as a roomer in a private home or lodging house or in permanent living quarters in a hotel or motel, but who is financially independent; or (3) two or more persons living together who use their income to make joint expenditures. Financial independence is determined by the three major expense categories: housing, food, and other living expenses. To be considered financially independent, at least two of the three major expense categories have to be provided entirely or in part by the respondent.

Reference person

The first member mentioned by the respondent when asked to "Start with the name of the person or one of the persons who owns or rents the home." It is with respect to this person that the relationship of other CU members is determined.

Income before taxes

The combined income earned by all CU members 14 years old or over during the 12 months preceding the interview. The components of income are: Wage and salary income, business income, farm income, Social Security income, Supplemental Security income, unemployment compensation, worker's compensation, public assistance, welfare, interest, dividends, pension income, income from roomers or boarders, other rental income, income from regular contributions, other income, and Food Stamps.

Income after taxes

Income before taxes minus personal taxes which includes Federal income taxes, state and local income taxes, and other taxes.

Complete income reporters

The distinction between complete and incomplete income reporters is based in general on whether the respondent provides values for major sources of income, such as wages and salaries, self-employment income, and social security income. Even complete income reporters may not provide a full accounting of all income from all sources. In the current survey, CUs that report across-the-board zero income are categorized as incomplete reporters.

Geographic regions

Data are presented for four major regions - Northeast, Midwest, South, and West. CUs are classified by region according to the address at which the CU was residing during the time of their participation in the survey. The regions comprise the following States:

Northeast - Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont.

Midwest - Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin.

South - Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia.

West - Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming.

XIII. APPENDIX 2 -- UNIVERSAL CLASSIFICATION CODE (UCC) TITLES

A. EXPENDITURE UCC's ON EXPN FILE

001000 Stocks, bonds, mutual funds
001100 Precious metals
001200 Miscellaneous investments

001400 Employment counseling & fees
 002000 Savings account deposit
 002100 Insurance other than health, hospital, vehicle and property
 002200 Retirement plans
 004000 Contributions
 004100 Cash gifts
 004190 Gifts not specified
 005000 Alimony and child support
 009000 Mortgage payment including coop
 009900 Property assessment
 010110 Flour
 010120 Prepared flour mixes
 010210 Cereal
 010310 Rice
 010320 Pasta, cornmeal, other cereal products
 020110 White bread
 020210 Bread other than white
 020310 Fresh biscuits, rolls, muffins
 020410 Cakes and cupcakes, fresh and other, excluding frozen
 020510 Cookies, excluding refrigerated dough
 020610 Crackers, excluding crumbs
 020620 Bread and cracker products
 020710 Doughnuts, sweet rolls, coffeecakes, fresh and other, excluding frozen
 020810 Frozen refrigerated and canned bakery products, such as biscuits, rolls, muffins, cakes, cupcakes, doughnuts, pies, tarts, turnovers, and miscellaneous products, including dough and batter
 020820 Pies, tarts, turnovers, fresh and other, excluding frozen
 030110 Ground beef, excluding canned
 030210 Chuck roast, excluding canned
 030310 Round roast, excluding canned
 030410 Other beef roast, excluding canned
 030510 Round steak, excluding canned
 030610 Sirloin steak, excluding canned
 030710 Other steak, excluding canned
 030810 Other beef, excluding canned
 040110 Bacon
 040210 Pork chops
 040310 Ham, excluding canned
 040410 Other pork, excluding canned
 040510 Pork sausage, excluding canned
 040610 Canned ham
 050110 Frankfurters, excluding canned
 050210 Bologna, liverwurst, salami, excluding canned
 050310 Other lunchmeat
 050410 Lamb and organ meats, excluding canned
 050900 Mutton, goat, game
 060110 Fresh and frozen whole chicken
 060210 Fresh or frozen chicken parts
 060310 Other poultry
 070110 Canned fish, seafood and shellfish
 070230 Fresh fish and shellfish
 070240 Frozen fish and shellfish
 080110 Eggs
 090110 Fresh milk all types
 090210 Cream
 100110 Butter

100210 Cheese
 100410 Ice cream and related products, including frozen yogurt
 100510 Other dairy products, including powdered milk, and fresh, canned and non-frozen yogurt
 110110 Apples
 110210 Bananas
 110310 Oranges
 110410 Other fresh fruits
 110510 Citrus fruits excluding oranges
 120110 Potatoes
 120210 Lettuce
 120310 Tomatoes
 120410 Other fresh vegetables
 130110 Frozen orange juice
 130121 Frozen fruits
 130122 Frozen fruit juices
 130211 Fresh fruit juices
 130212 Canned/bottled fruit juices
 130310 Canned fruits
 130320 Dried fruits
 140110 Frozen vegetables
 140210 Canned beans
 140220 Canned corn
 140230 Miscellaneous canned vegetables, not collected in a separate UCC
 140310 Other processed dried vegetables, such as squash, not collected in a separate UCC
 140320 Dried peas
 140330 Dried beans
 140340 Dried carrots, onions, leafy greens, and cabbage
 140410 Frozen vegetable juices
 140420 Fresh/canned vegetable juices
 150110 Candy and chewing gum
 150211 Sugar
 150212 Artificial sweeteners
 150310 Jams, jellies, preserves and other sweets
 160110 Margarine
 160211 Fats and oils
 160212 Salad dressings
 160310 Non-dairy cream substitutes
 160320 Peanut butter
 170110 Cola drinks
 170210 Other carbonated drinks
 170310 Coffee, roasted
 170410 Coffee, instant or freeze dried
 170510 Noncarbonated fruit flavored drinks, including lemonade-non frozen
 170520 Tea
 170530 Other noncarbonated beverages and ice, excluding coffee and tea
 180110 Soup
 180210 Frozen meals
 180220 Frozen prepared food other than meals
 180310 Potato chips and other snacks
 180320 Nuts
 180410 Salt, other seasonings & spices
 180420 Olives, pickles, relishes
 180510 Sauces and gravies
 180520 Other condiments
 180611 Prepared salads
 180612 Prepared desserts

180620 Baby food
 180710 Miscellaneous prepared foods including items such as canned meats (see UCC's 030110 - 030810, 040410 - 040510, 050110, 050310 - 050410, 060110 - 060310), fresh and canned ethnic foods, fresh and canned pizza
 190110 Lunch at restaurants, cafes, etc...
 190210 Dinner at restaurants, cafes, etc...
 190310 Snacks and non alcoholic beverages, including tip
 190320 Breakfast and brunch at restaurants, cafes, etc...
 190901 Food or board, at school and rooming/boarding houses
 190902 Catered affairs
 200111 Beer and ale at home
 200112 Nonalcoholic beer
 200210 Whiskey at home
 200310 Wine at home
 200410 Other alcoholic beverages at home
 200510 Beer and ale away from home
 200520 Wine away from home
 200530 Other alcoholic beverages away from home
 210110 Rent of dwelling, including deposit and parking fees
 210210 Lodging away from home
 210310 Housing for someone at school
 210900 Ground or land rent
 220000 Capital improvements, not specified
 220110 Fire/extended coverage insurance
 220120 Homeowners insurance
 220210 Property taxes
 220310 Contracted mortgage interest
 220400 Purchase of property or real estate
 220410 Home purchase
 220510 Capital improvements - commodities
 220610 Capital improvements - services
 220900 Parking, owned dwelling
 230000 Repair, maintenance, and improvements for built in dishwasher, garbage disposal, and range hood
 230110 Maintenance of property, including items such as ceiling repair, black top, brick, or masonry work, air conditioner repair, roof and awning repair, house painting, papering, chimney cleaning, electrical inspection, furnace inspection and repair, wiring, pest control, carpenter, plumber, etc...
 230120 Installed hard surface flooring
 230130 Installed wall-to-wall carpet
 230140 Repair disposal, dishwasher, range hood
 230900 Maintenance fees, such as service repair of property fees, management fees, homeowners association dues, condo fees, and community pool fees
 240110 Paint, wallpaper and supplies
 240120 Tools and equipment for painting and papering
 240210 Lumber, paneling, tile, awning, glass, plywood, doors, windows, screens, siding, roofing and fencing materials
 240220 Blacktop and masonry materials
 240310 Plumbing supplies, fixtures and equipment
 240320 Electric heating and air conditioning supplies and equipment
 240900 Soft surface floor covering
 250110 Fuel oil
 250210 Bottled or tank gas
 250220 Coal
 250900 Miscellaneous fuels, such as wood, kerosene, charcoal, oil mix for gas, lawnmower oil, lamp oil, duraflame log, and sterno

260110 Electricity
 260210 Utility - natural gas
 270000 Telephone service, including public pay phones
 270210 Water and sewerage maintenance
 270310 Community antenna or cable TV
 270410 Garbage, trash collection
 270510 Telephone interstate calls
 270610 Telephone intrastate calls
 270900 Septic tank cleaning
 270905 Steam heat
 280110 Bathroom linens
 280120 Bedroom linens
 280130 Kitchen and dining room linens
 280210 Curtains and drapes, excluding shower
 280220 Slipcovers, decorative pillows, and cushions
 280230 Sewing materials for slipcovers, curtains, and other home handiwork
 280900 Other linens
 290110 Mattress and springs
 290120 Other bedroom furniture
 290210 Sofas
 290310 Living room chairs
 290320 Living room tables
 290410 Kitchen and dining room furniture
 290420 Infants' furniture
 290430 Patio, porch or outdoor furniture
 290440 Modular wall units, shelves or cabinets, or other living room, family or rec-room furniture including desks
 300110 Refrigerator, home freezer
 300210 Washers
 300220 Dryers
 300310 Stoves, ovens
 300320 Microwave ovens
 300330 Portable dishwashers
 300410 Window air conditioners
 300900 Miscellaneous household appliances
 310110 Black and white TV's, and combination of TV with other items
 310120 Color TV console and combinations of TV with other items, such as TV with VCR
 310130 Color TV (portable and table models) and combinations of portable model color TV with other items, such as TV with radio
 310210 Video players, video recorders, video tape player, video tape recorder, video disc player, video camera receiver and recorder, and camcorder
 310220 Video cassettes, tapes and discs, laser discs, reels, prerecorded and blank video cassettes, video tapes, and diskettes
 310230 Video game cartridges, TV computer games and software, Atari cartridges and supplies, computer joystick, games, and game cartridges
 310311 Radio, not installed in vehicles
 310312 Phonograph or record player
 310313 Tape recorder and player
 310320 Sound components, component systems, amplifiers, receivers, turn tables, tape decks, tuners, stereos, speakers, and compact disc sound systems
 310331 Miscellaneous sound equipment
 310332 Sound equipment accessories
 A961 310334 Satellite dishes
 310340 Records, tapes, CD's, needles, styli, and record clubs
 310900 Accessories for electronic equipment
 320110 Room-size rugs and other non-permanent floor coverings

320120 Venetian blinds, window shades and other window coverings
 320130 Infants' equipment
 320140 Laundry and cleaning equipment
 320150 Outdoor equipment
 320210 Clocks
 320220 Lamps and other lighting fixtures
 320231 Other household decorative items, including fireplace equipment and accessories
 320232 Telephones and accessories
 320310 Plastic dinnerware
 320320 China and other dinnerware
 320330 Stainless, silver and other flatware
 320340 Glassware
 320350 Silver serving pieces
 320360 Serving pieces other than silver
 320370 Nonelectric cookware
 320380 Tableware, nonelectric kitchenware
 320410 Lawnmowing equipment and other yard machinery, powered and nonpowered
 320420 Power tools
 320430 Other hardware, including curtain and drapery hardware, rope, portable ladders, sheds, non-permanent shelves and shelving
 320511 Electric floor cleaning equipment
 320512 Sewing machines
 320521 Small electrical kitchen appliances
 320522 Portable heating and cooling equipment
 320610 Miscellaneous supplies and equipment, such as caulking compound, duct tape, carpet tape, carpet knife, bolts, screws, drill bits, door knobs, tool box, keys, mailbox, gutter screens, clamps, shelf brackets, tool table, work bench, etc...
 320620 Permanent hard surface floor covering
 320630 Landscaping items, such as grass, grass seed, trees, shrubs, plants, sod, and fork lift
 320901 Office furniture for home use
 320902 Non-powered tools
 320903 Fresh flowers or potted plants
 320904 Closet and storage items
 320905 Miscellaneous household equipment and parts
 320906 Electronic testing equipment
 330110 Soaps and detergents, excluding hand soaps
 330210 Other laundry and cleaning products
 330310 Paper towels, napkins, toilet tissue, facial tissue
 330410 Stationery, giftwrap and wrap accessories, greeting cards, pens, pencils, tape
 330510 Miscellaneous household products, including paper, plastic and foil products
 330610 Lawn and garden supplies, including outdoor plants
 340110 Postage
 340120 Delivery services
 340210 Babysitting or other home care for children
 340310 Housekeeping service, such as housekeeping, cooking, maid service, interior decorating, and carpet and upholstery cleaning services
 340410 Gardening and lawn care services, such as mowing, tree services, fertilizing, and yard work
 340510 Moving, storage, and freight express
 340520 Non-clothing household laundry or dry cleaning not coin operated
 340530 Non-clothing household laundry or dry cleaning - coin-operated
 340610 Repair of television, radio, and sound equipment, excluding installed in vehicles
 340620 Repair of household appliances; including stove, vacuum, washer, dryer, sewing machine, refrigerator, and calculator; excluding garbage disposal, range hood, and built-in dishwasher
 340630 Furniture repair, refinishing, or reupholstery
 340901 Rental or repair of lawnmowing equipment and other yard machinery, power and non-power

tools

340903 Miscellaneous home services and small repair jobs not already specified

340904 Rental of furniture

340906 Care for invalids, convalescents, handicapped or elderly persons in the CU

340907 Rental of household equipment items, such as refrigerators, home freezers, washers, microwave ovens, dishwashers, water cooler, stroller, china; excluding tools and lawn/garden equipment

340908 Rental of office equipment for non-business use, includes items such as calculators, typewriters, projectors, and other office machines.

340909 Rental of TV or radio sound equipment

340913 Repair and alterations of miscellaneous household equipment, furnishings, and textiles

350110 Tenants' insurance

360110 Men's suits

360120 Men's sportcoats and tailored jackets

360210 Men's coats, jackets, and furs

360311 Men's underwear

360312 Men's hosiery

360320 Men's sleepwear/loungewear

360330 Men's accessories

360340 Men's sweaters and vests

360350 Men's active sportswear

360410 Men's shirts

360511 Men's pants

360512 Men's shorts and shorts sets, excluding athletic

360901 Men's uniforms

370110 Boys' coats, jackets, and furs

370120 Boys' sweaters

370130 Boys' shirts

370211 Boys' underwear

370212 Boys' sleepwear/loungewear

370213 Boys' hosiery

370220 Boys' accessories

370311 Boys' suits, sportcoats, and vests

370312 Boys' pants

370313 Boys' shorts and shorts sets, excluding athletic

370901 Boys' uniforms and active sportswear

380110 Women's coats, jackets and furs

380210 Women's dresses

380311 Women's sportcoats and tailored jackets

380312 Women's vests, sweaters, and sweater sets

380313 Women's shirts, tops, and blouses

380320 Women's skirts and culottes

380331 Women's pants

380332 Women's shorts and shorts sets, excluding athletic

380340 Women's active sportswear

380410 Women's sleepwear/loungewear

380420 Women's undergarments

380430 Women's hosiery

380510 Women's suits

380901 Women's accessories

380902 Women's uniforms

390110 Girls' coats, jackets, and furs

390120 Girls' dresses and suits

390210 Girls' sport coats, tailored jackets, shirts, blouses, sweaters, sweater sets, and vests

390221 Girls' skirts, culottes, and pants

390222 Girls' shorts and shorts sets, excluding athletic

- 390230 Girls' active sportswear
- 390310 Girls' undergarments and sleepwear/loungewear
- 390321 Girls' hosiery
- 390322 Girls' accessories
- 390901 Girls' uniforms
- 400110 Men's footwear
- 400210 Boys' footwear
- 400220 Girls' footwear
- 400310 Women's footwear
- 410110 Infants' coats, jackets, and snowsuits
- 410120 Infants' rompers, dresses, and sweaters
- 410130 Infants' undergarments, including diapers
- 410140 Infants' sleeping garments
- 410901 Infants' accessories, hosiery, and footwear
- 420110 Sewing material for making clothes
- 420120 Sewing notions, patterns
- 430110 Watches
- 430120 Jewelry
- 430130 Travel items, including luggage, and luggage carriers
- 440110 Shoe repair and other shoe services
- 440120 Apparel laundry and dry cleaning - coin-operated
- 440130 Alteration, repair, tailoring of apparel and accessories
- 440140 Clothing rental
- 440150 Watch and jewelry repair
- 440210 Apparel laundry and dry cleaning not coin operated
- 440900 Clothing storage
- 450110 New cars
- 450210 New trucks, pick-ups, vans, or jeeps
- 450220 New motorcycles, motor scooters, or mopeds
- 450310 Lease payment (car lease)
- 450410 Lease payment (truck/pick-up/van/jeep lease)
- 460110 Used cars
- 460901 Used trucks or vans
- 460902 Used motorcycles, motor scooters, or mopeds
- 460903 Used aircraft
- 470111 Gasoline
- 470112 Diesel fuel
- 470114 Gasohol
- 470211 Motor oil
- 470220 Coolant/antifreeze, oil, brake & transmission fluids, additives, and radiator/cooling system protectant
- 480110 Tires (new, used or recapped); replacement and mounting of tires, and belting
- 480212 Vehicle products, such as wax, touch up paint, de-icer, protectant, polish, tar and bug remover, polish cloth, rubbing compound, auto freshner, etc...
- 480213 Battery replacement, floor mats, seat covers, filter, brake parts, and other equipment, supplies, parts, and accessories for auto; boating supplies and accessories
- 480214 Vehicle audio equipment, excluding labor
- 490000 Miscellaneous auto repair and servicing
- 490110 Body work, painting, repair and replacement of upholstery, vinyl/convertible top, and glass
- 490211 Clutch and transmission repair
- 490212 Drive shaft and rear-end repair
- 490220 Brake work, excluding brake adjustment
- 490231 Steering or front end repair
- 490232 Cooling system repair
- 490311 Motor tune-up
- 490312 Lubrication and oil changes

490313 Front end alignment, wheel balance and rotation
 490314 Shock absorber replacement
 490315 Brake adjustment
 490316 Gas tank repair and replacement
 490411 Exhaust system repair
 490412 Electrical system repair
 490413 Motor repair and replacement
 490900 Auto repair service policy
 500110 Vehicle insurance
 520111 Vehicle registration - state
 520112 Vehicle registration - local
 520310 Drivers' license
 520410 Vehicle inspection
 520511 Auto rental, excluding trips
 520521 Truck or van rental, excluding trips
 520531 Parking fees at garages, meters, and lots, excluding fees that are costs of property ownership in home city
 520541 Tolls
 520550 Towing charges
 520901 Docking and landing fees for boats and planes, boat ramp fees
 520902 Rental of motorcycle, motor scooters, moped, etc., including mileage charges
 520903 Rental of aircraft, including mileage charges
 520904 Rental of non camper-type trailer, such as for boat or cycle
 530110 Airline fares
 530210 Intercity bus fares
 530311 Intracity mass transit fares
 530412 Taxi fares
 530510 Intercity train fares
 530901 Ship fares
 530902 Private school bus
 530903 Car/van pool & non-motorized transportation
 540000 Prescription drugs and medicines
 550110 Purchase of eye glasses or contact lenses, excluding exam fee
 550210 Over-the-counter drugs
 550310 Topicals and dressings, such as band aids, guaze, cotton balls/rolls
 550320 Purchase of medical or surgical equipment for general use, such as thermometers, needles/syringes, ice bags, heating pads, (not including band aids, gauze, cotton rolls/balls)
 550330 Purchase of supportive or convalescent medical equipment, such as crutches, wheelchairs, braces, and ace bandages
 550340 Hearing aids
 550900 Recreational drugs
 560110 Physicians' services
 560210 Dental services
 560310 Eye exams, treatment or surgery, glass/lens service, glasses repaired
 560330 Lab tests and x-rays
 560400 Services by medical professionals other than physicians
 570000 Hospital care not specified
 570220 Care in convalescent in nursing home
 570230 Other medical care service, such as ambulance service
 570901 Rental of medical or surgical equipment for general use
 570902 Repair of medical equipment
 570903 Rental of supportive and convalescent equipment
 580000 Hospital and health insurance not spec.
 580110 Commercial health insurance
 580210 Blue Cross or Blue Shield

580310 Health maintenance plans
 580901 Medicare payments
 590110 Newspapers (single copy and subscriptions)
 590210 Magazines and periodicals (single copy and subscriptions)
 590220 Books purchased through book clubs
 590230 Books not purchased through book clubs
 590900 Newsletters
 600110 Outboard motor
 600120 Unpowered boats, trailers
 600130 Powered sports vehicles
 600210 Ping pong, pool tables, other similar items, general sports equipment, and health and exercise equipment
 600310 Bicycles
 600410 Camping equipment
 600420 Hunting and fishing equipment
 600430 Winter sports equipment
 600900 Water sports and miscellaneous sports equipment
 610110 Toys, games, hobbies, tricycles, and battery powered riders
 610120 Playground equipment
 610130 Musical instruments and accessories
 610210 Film
 610220 Other photographic supplies
 610230 Photographic equipment
 610310 Pet food
 610320 Pets, pet supplies and medicine for pets
 610901 Fireworks
 610902 Souvenirs
 610903 Visual goods
 620111 Membership fees for country clubs, health clubs, swimming pools tennis clubs, social or other recreational organizations, civic, service, or fraternal organizations
 620112 Membership fees for credit card memberships
 620113 Membership fees for automobile service clubs
 620121 Fees for participant sports, such as golf, tennis, and bowling
 620211 Admission fees for entertainment activities, including lectures, movie, theatre, concert, opera or other musical series
 620221 Admission fees to sporting events
 620310 Fees for recreational lessons or other instructions
 620320 Photographer fees
 620330 Film processing
 620410 Pet services
 620420 Veterinarian expenses for pets
 620510 Miscellaneous fees for admissions
 620610 Miscellaneous entertainment services
 620710 Camp fees
 620810 Rental and repair of sports, photographic and music equipment
 620911 Miscellaneous fees and pari-mutuel losses, licenses for sports and entertainment, passport fees, taxidermist fees
 620912 Rental of video cassettes, tapes, and discs
 620913 Coin-operated pinball/electronic video games
 620915 Passport fees
 630110 Cigarettes
 630210 Cigars, pipe tobacco, and other tobacco products
 630220 Smoking accessories
 630900 Marijuana
 640110 Hair care products
 640120 Non-electric articles for the hair

640130 Wigs, hairpieces, and toupees
 640210 Oral hygiene products, articles
 640220 Shaving needs
 640310 Cosmetics, perfume, cologne, bath preparations, hand soap, face and body powder, skin care products, nail preparations, manicure and eyemake-up implements and accessories
 640410 Deodorant, female hygiene products, miscellaneous personal care products and supplies
 640420 Electrical personal care appliances
 650110 Personal care services for females, including haircuts
 650210 Personal care services for males, including haircuts
 650900 Rental and repair of personal care appliances
 660000 School supplies., etc. - unspec., including reference books not in a set
 660110 School books, supplies, and equipment for college
 660210 School books, supplies, and equipment for elementary and high school
 660310 Encyclopedia and other sets of reference books
 660900 School books , supplies, and equipment for day care center, nursery school and other
 670110 Tuition for college
 670210 Tuition for elementary and high school
 670310 Other expenses for day care centers and nursery schools, including tuition
 670901 Tuition for other schools
 670902 Rentals of books and equipment, and other school-related expenses
 680110 Legal fees, excluding real estate closing costs
 680140 Funeral, burial or cremation expenses
 680210 Safe deposit box rental
 680220 Charges for checking accounts and other banking services, excluding safe deposit
 680901 Purchase and upkeep of cemetery lots or vaults
 680902 Accounting fees
 680903 Miscellaneous personal services, advertising, fines, duplicating services
 690110 Computers for non-business use, hardware and software excluding video games
 690114 Computer information services
 690210 Telephone answering devices
 690220 Calculators
 690230 Typewriters and other office machines for non-business use
 999000 Home ownership expense not specified
 999900 Taxes not specified
 999912 Unidentifiable items - Parts 1 and 2
 999935 Unidentifiable items - Parts 3, 4, and 5

B. INCOME AND RELATED UCC's ON DTAB FILE

*L denotes UCC's could have negative values

800700 Meals received as pay
 800710 Rent received as pay
 800910 Payroll deductions for government retirement
 800920 Payroll deductions for railroad retirement
 800931 Payroll deductions for private pensions
 800932 Non-payroll deposit to individual retirement plan, such as IRA's
 800940 Payroll deductions for social security
 900000 Wages and salaries
 *L 900010 Net business income
 *L 900020 Net farm income

900030	Social security and railroad retirement income
900040	Pensions and annuities
900050	Dividends, royalties, estates, or trusts
*L 900060	Income from roomers and boarders
*L 900070	Other rental income
900080	Interest from saving accounts or bonds
900090	Supplemental security income
900100	Unemployment compensation
900110	Worker's compensation and veterans payments including education benefits
900120	Public assistance or welfare including money received from job training grants such as job corps
900131	Child support payments received
900132	Other regular contributions received including alimony
900140	Other income including money received from care of foster children, cash scholarships and fellowships or stipends not based on working
900150	Food stamps
910000	Lump sum payments from estates, trusts, royalties, alimony, child support, prizes or games of chance, or from persons outside of the CU
910010	Money from sale of household furnishings, equipment, clothing, jewelry, pets or other belongings, excluding the sale of vehicles or property
910020	Overpayment on social security
910030	Refund from insurance policies
910040	Refunds from property taxes
910041	Lump sum child support payments received
950000	Federal income tax
*L 950001	Federal income tax refunds
950010	State and local income tax
*L 950011	State and local income tax refunds
950021	Other taxes
950022	Personal property taxes
*L 950023	Other tax refunds
*L 980000	Income before taxes
980010	Family size
980020	Age of reference person
980030	Number of earners
980040	Number of vehicles
980050	Number of persons under 18
980060	Number of persons 65 and over
*L 980070	Income after taxes

The following UCC's contain values of either 100 or 0 depending on whether the CU satisfies the condition. For example, if the CU owns the home, then UCC 980090, homeowner, will have a value of 100, if not, the UCC will have a value of zero. These UCC's are used at BLS to compute percentages for the published tables.

980090	Percent homeowner
980210	Percent male reference person
980220	Percent female reference person
980230	Percent homeowner with mortgage
980240	Percent homeowner without mortgage
980250	Percent homeowner with mortgage not reported
980260	Percent renter
980270	Percent black reference person
980280	Percent non-black reference person

980290	Percent reference person with elementary education
980300	Percent reference person with high school education
980310	Percent reference person with college education
980320	Percent reference person with no education and other
980330	Percent vehicle owner

XIV. APPENDIX 3 -- UCC AGGREGATION

The following shows the UCC aggregation used in the sample program. This information is provided on the AGGregation and LABel files (Section III.E.5.)

Food	010110-190902, 200112
Food at home	010110-180710, 200112
Cereal and cereal products	010110-010320
Bakery products	020110-020820
Beef	030110-030810
Pork	040110-040610
Other meats	050110-050900
Poultry	060110-060310
Fish and seafood	070110-070240
Eggs	080110
Fresh milk and cream	090110-090210
Other dairy products	100110-100510
Fresh fruits	110110-110510
Fresh vegetables	120110-120410
Processed fruits	130110-130320
Processed vegetables	140110-140420
Sugar and other sweets	150110-150310
Fats and oils	160110-160320
Miscellaneous foods	180110-180710
Nonalcoholic beverages	170110-170530, 200112
Food away from home	190110-190902
Alcoholic beverages	200111, 200210-200530
Housing	250110-300900, 320110-340913
Fuel and utilities	250110-270000
Household operations	340520, 340913
Housekeeping supplies	330110-340120
Household furnishings and equipment	300900-320905
Apparel and services	360110-360901, 370110-370901, 380110-380902, 390110-390901, 410110-410140, 410901, 400110- 400310, 420110-430120, 440110-440900
Men, 16 and over	360210-360901
Boys, 2 to 15	370130-370220
Women, 16 and over	380110-380902
Girls, 2 to 15	390120-390322
Children under 2	410120-410901
Footwear	400110-400310
Other apparel products and services	420110-430120, 440120-440210
Transportation	470111-470220, 480212-49000, 49316, 520541
Non-prescription drugs and supplies	550210-550410, 570902
Entertainment	310311-310332, 310334, 600210-620913

Radios, sound equipment	310311-310332, 310334
Pet food and supplies	610310-610320
Toys, games, playground equipment	610110-610120
Other entertainment supplies, equipment	600210-600420, 610220-610903, 620320, 620913
Personal care products and services	640110-640410
Miscellaneous	620911, 630220, 680903

XV. APPENDIX 4 -- FMLY AND MEMB VARIABLES ORDERED BY START POSITION

This appendix lists FMLY and MEMB variables in the order that they appear on the files. Sections III.E.1. and III.E.2. contain detailed descriptions of these variables arranged on a functional basis.

A. FMLY FILE

Variable	Start Position	Variable	Start Position	Variable	Start Position
NEWID	1	FAM_IZE	80	FS_AMT1_	212
ADDFEDX	9	FAM_TYPE	81	FS_AMT2	213
ADDFEDX_	17	FAM__YPE	82	FS_AMT2_	221
ADDOTHX	18	FBSNSX	83	FS_AMT3	222
ADDOTHX_	26	FBSNSX_	91	FS_AMT3_	230
ADDSTAX	27	FD_STMPS	92	FS_AMT4	231
ADDSTAX_	35	FD_S_MPS	93	FS_AMT4_	239
AGE_REF	36	FEDREFX	94	FS_AMT5	240
AGE_REF_	38	FEDREFX_	102	FS_AMT5_	248
AGE2	39	FFARMX	103	FS_AMT6	249
AGE2_	41	FFARMX_	111	FS_AMT6_	257
BLS_URBN	42	FFEDTXX	112	FS_AMT7	258
CUTENURE	43	FFEDTXX_	120	FS_AMT7_	266
CUTE_URE	44	FGVX	121	FS_AMT8	267
DESCRIP	45	FGVX_	129	FS_AMT8_	275
DESCRIP_	47	FINCAFTX	130	FS_DATE1	276
DIVX	48	FINC_FTX	138	FS_D_TE1	284
DIVX_	56	FINCBEFX	139	FS_DATE2	285
EARNCOMP	57	FINC_EFX	147	FS_D_TE2	293
EARN_OMP	58	FINLWT21	148	FS_DATE3	294
EARNX	59	FIRAX	159	FS_D_TE3	302
EARNX_	67	FIRAX_	167	FS_DATE4	303
EDUC_REF	68	FJSSDEDX	168	FS_D_TE4	311
EDUC0REF	70	FJSS_EDX	176	FS_DATE5	312
EDUCA2	71	FPVTX	177	FS_D_TE5	320
EDUCA2_	73	FPVTX_	185	FS_DATE6	321
EMPLTYP1	74	FREEMLX	186	FS_D_TE6	329
EMPL_YP1	75	FREEMLX_	194	FS_DATE7	330
EMPLTYP2	76	FRRX	195	FS_D_TE7	338
EMPL_YP2	77	FRRX_	203	FS_DATE8	339
FAM_SIZE	78	FS_AMT1	204	FS_D_TE8	347

Variable	Start Position	Variable	Start Position	Variable	Start Position
FS_MTHI	348	OTHREFX_	525	WEEKI	656
FS_MTHI_	350	OTHRNTX	526	WEEKI_	657
FSS_RRX	351	OTHRNTX_	534	WEEKN	658
FSS_RRX_	359	PENSIONX	535	WELFRX	659
FSTATXX	360	PENS_ONX	543	WELFRX_	667
FSTATXX_	368	PERSLT18	544	WHYNWRK1	668
FSUPPX	369	PERS_T18	546	WHYN_RK1	669
FSUPPX_	377	PERSOT64	547	WHYNWRK2	670
FWAGEX	378	PERS_T64	549	WHYN_RK2	671
FWAGEX_	386	PERSTAX	550	WK_WRKD1	672
HRSPRWK1	387	PERSTAX_	558	WK_W_KD1	674
HRSP_WK1	390	PICK_UP	559	WK_WRKD2	675
HRSPRWK2	391	OCCULIS1	561	WK_W_KD2	677
HRSP_WK2	394	OCCU_IS1	563	WRKRSX	678
INC_RNKU	395	POPSIZE	564	WRKRSX_	686
INC__NKU	404	PTAXREFX	565	WTREP01	687
INSREFX	405	PTAX_EFX	573	WTREP02	698
INSREFX_	413	RACE2	574	WTREP03	709
INTX	414	RACE2_	575	WTREP04	720
INTX_	422	REC_FS	576	WTREP05	731
JFS_AMT	423	REC_FS_	577	WTREP06	742
JFS_AMT_	431	REF_RACE	578	WTREP07	753
JGRCFDMV	432	REF__ACE	579	WTREP08	764
JGRC_DMV	438	REGION	580	WTREP09	775
JGRCFDWK	439	REGION_	581	WTREP10	786
JGRC_DWK	445	RESPSTAT	582	WTREP11	797
JGROCYMV	446	RESP_TAT	583	WTREP12	808
JGRO_YMV	452	ROOMX	584	WTREP13	819
JGROCYWK	453	ROOMX_	592	WTREP14	830
JGRO_YWK	459	SALEX	593	WTREP15	841
LUMPX	460	SALEX_	601	WTREP16	852
LUMPX_	468	SEX_REF	602	WTREP17	863
MARITAL1	469	SEX_REF_	603	WTREP18	874
MARI_AL1	470	SEX2	604	WTREP19	885
NO_EARNR	471	SEX2_	605	WTREP20	896
NO_E_RNR	473	SMSASTAT	606	WTREP21	907
NONERNX	474	SSREFX	607	WTREP22	918
NONERNX_	482	SSREFX_	615	WTREP23	929
OCCEXPNX	483	STATREFX	616	WTREP24	940
OCCE_PNX	491	STAT_EFX	624	WTREP25	951
OCCULIS2	492	STRTDAY	625	WTREP26	962
OCCU_IS2	494	STRTMNTH	627	WTREP27	973
ORIGIN1	495	STRTYEAR	629	WTREP28	984
ORIGIN1_	496	TAXPROPX	633	WTREP29	995
ORIGIN2	497	TAXP_OPX	641	WTREP30	1006
ORIGIN2_	498	TYPOWND	642	WTREP31	1017
OTHINX	499	TYPOWND_	643	WTREP32	1028
OTHINX_	507	UNEMPX	644	WTREP33	1039
OTHRECX	508	UNEMPX_	652	WTREP34	1050
OTHRECX_	516	VEHQ	653	WTREP35	1061
OTHREFX	517	VEHQ_	655	WTREP36	1072

Variable	Start Position	Variable	Start Position	Variable	Start Position
WTREP37	1083	OTHDAIRY	1303	HHID	1510
WTREP38	1094	FRSHFRUT	1315	HHID_	1513
WTREP39	1105	FRSHVEG	1327	CHILDAge	1514
WTREP40	1116	PROCFRUT	1339	CHIL_Age	1515
WTREP41	1127	PROCVeG	1351	INCLASS	1516
WTREP42	1138	SWEETS	1363	STATE	1518
WTREP43	1149	NONALBEV	1375	STATE_	1520
WTREP44	1160	OILS	1387	CHDOThX	1521
FOODTOT	1171	MISCFOOD	1399	CHDOThX_	1529
FOODHOME	1183	FOODAWAY	1411	ALIOThX	1530
CEREAL	1195	ALCBEV	1423	ALIOThX_	1538
BAKEPROD	1207	SMOKSUPP	1435	CHDLMPX	1539
BEEF	1219	PET_FOOD	1447	CHDLMPX_	1547
PORK	1231	PERSPROD	1459	POVERTY	1548
OTHMEAT	1243	PERSSERV	1471	POVERTY_	1549
POULTRY	1255	DRUGSUPP	1483		
SEAFOOD	1267	HOUSKEEP	1495		
EGGS	1279	HH_CU_Q	1507		
MILKPROD	1291	HH_CU_Q_	1509		

B. MEMB FILE

Variable	Start Position	Variable	Start Position	Variable	Start Position
NEWID	1	EMPL_YPE	76	RACE	151
AGE	9	FARMX	77	RACE_	152
AGE_	11	FARMX_	85	RRX	153
ANFEDTXX	12	FEDTXX	86	RRX_	161
ANFE_TXX	20	FEDTXX_	94	SCHLNCHQ	162
ANGVX	21	GROSPAYX	95	SCHL_CHQ	164
ANGVX_	29	GROS_AYX	103	SCHLNCHX	165
ANPVTX	30	GVX	104	SCHL_CHX	173
ANPVTX_	38	GVX_	112	SEX	174
ANRRX	39	HRSPERWK	113	SEX_	175
ANRRX_	47	HRSP_RWK	116	SLFEMPSS	176
ANSTATXX	48	IRAX	117	SLFE_PSS	182
ANST_TXX	56	IRAX_	125	SS_RRX	183
ANYRAIL	57	JSSDEDX	126	SS_RRX_	191
ANYRAIL_	58	JSSDEDX_	132	STA_SUPP	192
ANYSSINC	59	MARITAL	133	STA_UPP	193
ANYS_INC	60	MARITAL_	134	STATXX	194
BSNSX	61	MEMBNO	135	STATXX_	202
BSNSX_	69	OCCULIST	137	SUPPX	203
CU_CODE1	70	OCCU_IST	139	SUPPX_	211
CU_C_DE1	71	ORIGIN	140	US_SUPP	212
EDUCA	72	ORIGIN_	141	US_SUPP_	213
EDUCA_	74	PVTX	142	WAGEX	214
EMPLTYPE	75	PVTX_	150	WAGEX_	222

Variable	Start Position	Variable	Start Position	Variable	Start Position
WHYNOWRK	223	SS_RRQ_	232	IN_COLL	244
WHYN_WRK	224	SOCRRX	233	IN_COLL_	245
WKS_WRKD	225	SOCRRX_	241	MEDICARE	246
WKS__RKD	227	ARM_FORC	242	MEDI_ARE	247
SS_RRQ	228	ARM__ORC	243		

XVI. APPENDIX 5--PUBLICATIONS AND DATA RELEASES

A list of publications containing data from the Consumer Expenditure Survey program appears below. Bulletins may be purchased from the Chicago regional sales center, from the U.S. Government Printing Office, Washington D.C., 20402, or from National Technical Information Service, U.S. Department of Commerce, Springfield, Virginia 22161. To place a telephone order with National Technical Information Service, call (703)-487-4650, or for a rush order, call 1(800)-553-NTIS.

Consumer Expenditure Survey, 1996-97, Bulletin (expected release Summer 1999)	Consumer Unit income and expenditures, integrated data from Interview and Diary Surveys, classified by consumer unit characteristics: one way and cross tabulations, relative and aggregate shares. 64 tables.
Consumer Expenditures in 1997, Report 927 (1999)	Consumer unit income and expenditures, integrated data from Diary and Interview Surveys, classified by consumer unit characteristics. 10 tables. Available on request (202)-606-6900.
Consumer Expenditures in 1996, Report 926 (1998)	Consumer unit income and expenditures, integrated data from Diary and Interview Surveys, classified by consumer unit characteristics. 10 tables. Available on request (202)-606-6900.
Consumer Expenditure Survey, 1994-95, Bulletin 2492 (1997)	Consumer Unit income and expenditures, integrated data from Interview and Diary Surveys, classified by consumer unit characteristics: one way and cross tabulations, relative and aggregate shares. 64 tables.
Consumer Expenditures in 1995, Report 911 (1997)	Consumer unit income and expenditures, integrated data from Diary and Interview Surveys, classified by consumer unit characteristics. 10 tables. Available on request (202)-606-6900.
Consumer Expenditure Survey, 1992-93, Bulletin 2462 (1995)	Consumer unit income and expenditures, integrated data from Diary and Interview Surveys, classified by consumer unit characteristics: one way and cross tabulations, relative and aggregate shares. 60 tables, 245 pages. Available at the Government Printing Office, stock number 029-001-03214-5, \$15.00.
Consumer Expenditure Survey, 1990-91, Bulletin 2425 (1993)	Consumer unit income and expenditures, integrated data from Diary and Interview Surveys, classified by consumer unit characteristics: one way and cross tabulations, relative and aggregate shares. 60 tables, 256 pages. NTIS Accession No. PB95-190948, \$36.50 for paper copy, \$17.50 for microfiche.

Consumer Expenditure Survey, 1988-89, Bulletin 2383 (1991)	Consumer unit income and expenditures, integrated data from Interview and Diary Surveys, classified by consumer unit characteristics: one way and cross tabulations. 40 tables, 199 pages. NTIS Accession #PB92130061, \$36.00 for paper copy, \$17.50 for microfiche.
Consumer Expenditure Survey, 1987, Bulletin 2354 (1990)	Consumer unit income and expenditures, integrated data from Interview and Diary Surveys, classified by consumer unit characteristics; one way and cross tabulations. 29 tables, 153 pages. NTIS Accession #PB92131622, \$27.00 for paper copy, \$12.50 for microfiche.
Consumer Expenditure Survey: Integrated Survey Data, 1984-86, Bulletin 2333 (1989)	Consumer unit income and expenditures, integrated data from Interview and Diary Surveys, classified by consumer unit characteristics; one way and cross tabulation. 34 tables, 171 pages. NTIS Accession #PB92131515, \$27.00 for paper copy, \$12.50 for microfiche.
Consumer Expenditure Interview Survey: Quarterly Data, 1984-1987, Bulletin 2332 (1989)	Consumer unit income and expenditures from the Interview Survey presented by quarter, classified by consumer unit characteristics; region, size, age, quintiles, income before taxes, and tenure tables included. 100 tables, 113 pages. NTIS Accession #PB92131523, \$27.00 for paper copy, \$12.50 for microfiche.

CONSUMER EXPENDITURE SURVEY: QUARTERLY DATA FROM THE INTERVIEW SURVEY

These quarterly reports present selected expenditure data and include a brief analysis of trends in consumer spending or other topics related to the Consumer Expenditure Survey. Requests for these reports can be made at (202) 606-6900.

CONSUMER EXPENDITURE DATA ON THE INTERNET

Integrated Diary and Interview data from 1984 forward can now be accessed on the Internet. The Internet address for these data is STATS.BLS.GOV. Tables containing average annual expenditures for major expenditure categories and by various consumer unit characteristics may be found under CE, within the SPECIAL.REQUESTS directory. World Wide Web users can access these same data at [HTTP://STATS.BLS.GOV](http://STATS.BLS.GOV) and then selecting, in order, hypertext for -Surveys and Programs-, -Prices and Living Conditions-, and -Consumer Expenditure Survey-.

FAX ON DEMAND - FAXSTAT

FAXSTAT contains information and data that may be faxed to users from a touch-tone phone 24 hours a day -- 7 days a week. To receive FAXSTAT transmissions dial (202) 606-6FAX and follow the voice prompts. Consumer Expenditure Survey data that are accessible on FAXSTAT are for the most recent year available

PUBLIC-USE TAPES

Public-use tapes for the Diary and Interview Surveys are now available for single years from 1984 to 1995, and as two-year tapes for 1982-83. Seven public-use tapes are available from the 1972-73 survey including Diary Survey, - Detailed Food Quantity tapes; and integrated adjusted Quarterly

Interview Survey - Summary, Detailed, Consumer Durables, and Clothing and Household Textiles tapes. Information about the tapes is available from the BLS national office.

COMPACT DISKS

Consumer Expenditure Survey microdata on compact disk are available from the Bureau of Labor Statistics for 1972-73, 1984-85, 1990-91, 1992-93, 1994, 1995, and 1996. The 1984-85 through 1996 releases contain Interview and Diary data, while the 1972-73 CD includes Interview data only. The 1984-85, and the 1990 files (of the 1990-91 CD) include selected EXPN data, while the 1991 files (from the 1990-91 CD) and the 1992-93 compact disk do not. In addition to the Interview and Diary data, the 1994, 1995, and 1996 CD's include the complete collection of EXPN files. A 1984-94 "multi-year" compact disk that presents Interview FMLY file data is also available. In addition to the microdata, the CD's also contain the same integrated Diary and Interview tabulated data that are found on the Consumer Expenditure Survey data diskettes. (see below)

DISKETTES

Diskettes containing integrated Diary and Interview survey data on consumer expenditures, income, and characteristics are available for the years 1984 through 1996. The diskettes are for use with IBM-compatible microcomputers with 3 1/2" disk drives. Users may specify either a Lotus 1-2-3 or an ASCII format.

The data on the diskettes are average annual expenditures by American consumers. They are presented in tables classified by 12 standard characteristics: quintiles of income, income class, age, size of consumer unit, composition of consumer unit, number of earners, housing tenure, race, type of area (urban-rural), region, occupation, and origin. Also on these diskettes are: data classified by income before taxes, cross-tabulated by age, by family size, or by region; data for selected Metropolitan Statistical Areas,; and data for single persons classified by gender, cross-tabulated by age or by income. Expenditure categories in these tables are similar to those shown in the tables of the bulletin publications. For a more detailed description and an order form contact the BLS national office.

XVII. INQUIRIES, SUGGESTIONS, AND COMMENTS

If you have any questions, suggestions, or comments about the survey, the microdata, or its documentation, please call.(202) 606-6900

Written suggestions and comments should be forwarded to:

Division of Consumer Expenditure Surveys
Branch of Information and Analysis
Bureau of Labor Statistics, Room 3985
2 Massachusetts Ave. N.E.
Washington, DC. 20212-0001

The Bureau of Labor Statistics will use these responses in planning future releases of the microdata files.