2010 CONSUMER EXPENDITURE DIARY SURVEY
PUBLIC USE MICRODATA
User's Documentation
September 27, 2011

U.S. Department of Labor
Bureau of Labor Statistics
Division of Consumer Expenditure Survey

TABLE OF CONTENTS

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 or CUs) of average expenditures in news releases, reports, and articles in the Monthly Labor Review. Tabulated CE data are also available on the Internet and by facsimile transmission (see Section XVI. Appendix 5). The microdata are available on CD-ROMs.

These microdata files present detailed expenditure and income data for the Diary component of the CE for 2010. They include weekly expenditure (EXPN), annual income (DTAB) files, and imputed income files (DTID). The data in EXPN, DTAB, and DTID files are categorized by a Universal Classification Code (UCC). The advantage of the EXPN and DTAB files is that with the data classified in a standardized format, the user may perform comparative expenditure (income) analysis with relative ease. The FMLY and MEMB files present data on the characteristics and demographics of CUs and CU members. The summary level expenditure and income information on the FMLY files permits the data user to link consumer spending, by general expenditure category, and household characteristics and demographics on one set of files.

Estimates of average expenditures in 2010 from the Diary survey, integrated with data from the Interview survey, are published in *Consumer Expenditures in 2010.* 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, 2010”.

II. CHANGES FROM THE 2009 MICRODATA FILES

A. FMLY Files

No changes in 2010

B. MEMB Files

No changes in 2010

C. EXPN Files

No changes in 2010

D. DTAB Files

No changes in 2010

E. DTID Files

No changes in 2010

III. FILE INFORMATION

The microdata on the CD-ROM are provided as SAS, STATA, SPSS data sets or ASCII text and comma-delimited files. The 2010 Diary release contains five sets of data files (FMLY, MEMB, EXPN, DTAB, DTID) and three processing files. The FMLY, MEMB, EXPN, DTAB, and DTID files are organized by the quarter of the calendar year in which the data were collected. There are four quarterly data sets for each of these files. The FMLY files contain CU characteristics, income, and summary level expenditures; the MEMB files contain member characteristics and income data; the EXPN files contain detailed weekly expenditures at the UCC level; the DTAB files contains the CU's reported income values or the mean of the five imputed income values in the multiple imputation method; and the DTID files contain the five imputed income values.

The three processing files enhance computer processing and tabulation of data, and provide descriptive information on item codes. The three processing files are: an aggregation scheme file used in the published consumer expenditure tables (DSTUB), a UCC file that contains UCCs and their abbreviated titles, identifying the expenditure, income, or demographic item represented by each UCC, and a sample program file that contains the computer program used in Section VII.A. SAMPLE PROGRAM of the documentation. The processing files are further explained in Section III.E.5. PROCESSING FILES.

In addition to these processing files, there is a "User's Guide to Income Imputation in the CE", which includes information on how to appropriately use the imputed income data.

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

A. DATA SET NAMES

The file naming convention in the ASCII subfolder (X:\Data\2010_ASCII\diary10) is as follows: (where "X" references the designated drive for your CD)

```
\DIARY10\FMLYD101.txt   (Diary FMLY file for first quarter, 2010)
\DIARY10\MEMBD101.txt   (Diary MEMB file for first quarter, 2010)
\DIARY10\EXPND101.txt   (Diary EXPN file for first quarter, 2010)
\DIARY10\DTABD101.txt   (Diary DTAB file for first quarter, 2010)
\DIARY10\DTBID101.txt   (Diary IMPUTED DTAB file for, 2010 Q1)
\DIARY10\FMLYD102.txt   (etc.)
\DIARY10\MEMBD102.txt
\DIARY10\EXPND102.txt
\DIARY10\DTABD102.txt
\DIARY10\DTBID102.txt
\DIARY10\FMLYD103.txt
\DIARY10\MEMBD103.txt
\DIARY10\EXPND103.txt
\DIARY10\DTABD103.txt
\DIARY10\DTBID103.txt
\DIARY10\FMLYD104.txt
\DIARY10\MEMBD104.txt
\DIARY10\EXPND104.txt
\DIARY10\DTABD104.txt
\DIARY10\DTBID104.txt
\DIARY10\UCCD10.txt
```

The file naming convention in the SAS subfolder X:\\Data\2010_SAS\diary10 is listed in the table below. The STATA, ASCII comma-delimited, and SPSS files use the same dataset names as SAS, but have a different file extension as follows:
Comma-delimited ASCII files: *.csv
STATA files: *.dta
SPSS files: *.sav

```
\DIARY10\FMLD101.sas7bdat     (Diary FMLY file for first quarter, 2010)
\DIARY10\MEMD101.sas7bdat     (Diary MEMB file for first quarter, 2010)
\DIARY10\EXPD101.sas7bdat     (Diary EXPN file for first quarter, 2010)
\DIARY10\DTBD101.sas7bdat     (Diary DTAB file for first quarter, 2010)
\DIARY10\DTID101.sas7bdat     (Diary Imputed DTAB file for, 2010 Q1)
\DIARY10\FMLD102.sas7bdat
\DIARY10\MEMD102.sas7bdat
\DIARY10\EXPD102.sas7bdat
\DIARY10\DTBD102.sas7bdat
\DIARY10\DTID102.sas7bdat
\DIARY10\FMLD103.sas7bdat
\DIARY10\MEMD103.sas7bdat
\DIARY10\EXPD103.sas7bdat
\DIARY10\DTBD103.sas7bdat
\DIARY10\DTID103.sas7bdat
\DIARY10\FMLD104.sas7bdat
\DIARY10\MEMD104.sas7bdat
\DIARY10\EXPD104.sas7bdat
\DIARY10\DTBD104.sas7bdat
\DIARY10\DTID104.sas7bdat
\DIARY10\UCCD10.txt
```

Note: All data files are compressed so the data can fit on one CD-ROM. These files can be uncompressed using most unzip utilities.

B. RECORD COUNTS AND LOGICAL RECORD LENGTHS PER QUARTER

The following are number of records and the logical record lengths (LRECL) in each data set. The OBS count is also applicable to the STATA and SPSS files:

ASCII data set	SAS data set	2010 Record Count	2010 LRECL
FMLYD101.txt	FMLD101.sas7bdat	3725	3119
MEMBD101.txt	MEMD101.sas7bdat	9298	775
EXPND101.txt	EXPD101.sas7bdat	130295	40
DTABD101.txt	DTBD101.sas7bdat	63052	28
DTBID101.txt	DTID101.sas7bdat	95631	29
FMLYD102.txt	FMLD102.sas7bdat	3619	3119
MEMBD102.txt	MEMD102.sas7bdat	9257	775
EXPND102.txt	EXPD102.sas7bdat	130265	40
DTABD102.txt	DTBD102.sas7bdat	61362	28
DTBID102.txt	DTID102.sas7bdat	92245	29
FMLYD103.txt	FMLD103.sas7bdat	3511	3119
MEMBD103.txt	MEMD103.sas7bdat	8592	775
EXPND103.txt	EXPD103.sas7bdat	119867	40
DTABD103.txt	DTBD103.sas7bdat	59099	28
DTBID103.txt	DTID103.sas7bdat	88978	29

ASCII data set	SAS data set	2010 Record Count	2010 LRECL
FMLYD104.txt	FMLD104.sas7bdat	3441	3119
MEMBD104.txt	MEMD104.sas7bdat	8518	775
EXPND104.txt	EXPD104.sas7bdat	119285	40
DTABD104.txt	DTBD104.sas7bdat	63110	28
DTBID104.txt	DTID104.sas7bdat	95030	29

C. DATA FLAGS:

Data fields on the FMLY and MEMB files are explained by flag variables following the data field. The names of the flag variables 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, so that PENSIONX becomes PENS_ONX, 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 data field contains a valid or good data value.

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

Some Primary Sampling Units (PSUs) in some states are given "false" STATE codes for nondisclosure reasons. See Section IV.A.CU CHARACTERISTICS AND INCOME FILE (FMLY) on topcoding of CU characteristics and income for more detail.

D. INCOME IMPUTATION

Beginning in 2004, the CE has implemented multiple imputation of income data. Imputation allows income values to be estimated when they are not reported. Many income variables and other income related variables will be imputed using a multiple imputation process. These imputed income values will be included in the FMLY, MEMB, DTAB, and DTID files. The multiple imputation process derives five imputation values, and a mean imputation value, per selected income variable. More information on the imputation process and how to appropriately use the data are found in the document "User's guide to Income Imputation in the CE".

In the public-use microdata, not all of the imputed income variables will contain the derived imputation values. For some income variables, the five derived imputations are excluded and only the mean of those imputations is available. For these variables, there are 3 associated income variables in the FMLY and MEMB files (INCOMEM, INCOMEM_, and INCOMEI). For all other imputed income variables, there are 7 associated variables in the FMLY and MEMB files:

INCOME1 - the first imputed income value or the reported income value, if non-missing
INCOME2 - the second imputed income value or the reported income value, if non-missing
INCOME3 - the third imputed income value or the reported income value, if non-missing
INCOME4 - the fourth imputed income value or the reported income value, if non-missing
INCOME5 - the fifth imputed income value or the reported income value, if non-missing
INCOMEM - the mean of the five imputed income values
INCOMEM_ - the flag variable for the imputed variable (see section III.C. Data Flags)
INCOMEI - the imputation indicator

Income variables that have imputed values as components (ex: FINCBEFM) will also have 5 imputed values and a mean based on each of the imputed components.

The imputation indicator variable is a 3 digit number that is coded as follows:

The first digit in the 3 digit code defines the imputation method. The meanings are:

1: No Imputation
2: Multiple Imputation due to invalid blank only
3: Multiple Imputation due to bracketing only
4: Multiple Imputation due to invalid blanks and bracketing
5: Multiple Imputation due to conversion of a valid blank to an invalid blank (this occurs only when initial values for all sources of income for the CU were valid blanks).

The meaning of the last two digits of the three digit code differs depending on whether you are looking at one of the components of overall income, like FWAGEXM, or you are looking at the summary level variable FINCBEFM. For the components the last 2 digits represent the number of family members who had their data imputed for that source. For example, if a family had a value of 302 for FWAGEXI that would mean that 2 of the members in the family had their salary income imputed and that in both cases the imputation was due to bracketing only. For the summary level variable FINCBEFM which is a summation of all of the income components, the last 2 digits represent the number of income sources imputed for each member all added together. So, for example, if a family had 3 members and 2 had salary income imputed due to invalid blank only, and 2 had nonfarm income imputed due to bracketing only, and that was the only income data imputed for members of that family, then FWAGEXI for the family would be 202, FBSNSXI would be 302, and FINCBEFI would be 404.

The DTAB file includes income UCCs mapped from the associated INCOMEM variables and the income variables that are not imputed in the FMLY files. The DTID file includes UCCs mapped from income variables subject to income imputation, including the variable IMPNUM to indicate the imputation number 1 - 5.

E. 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.

Data fields for variables on the microdata files have either numeric or character values. The format column in the diary data dictionary distinguishes whether a variable is numeric (NUM) or character (CHAR) and shows the number of field positions the variable occupies. Variables that 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.

In addition to format, the diary data dictionary gives an item description, questionnaire source, identification of codes where applicable, and start position for each variable.

A star (*) is shown in front of new variables, those which have changed in format or definition, and those which have been deleted.

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

*D(Yxxq) identifies a variable that 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(Y101) indicates the variable is deleted starting with the data file of the first quarter of 2010.

*N(Yxxq) identifies a variable that 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.

*C(Yxxq) identifies a variable whose description has been changed. The year and quarter are identified by the ‘xx’ and ‘q’ for new variables in the same way as for new and deleted variables.

*L indicates that the variable can contain negative values.

F. NOTES ON FILES

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, and 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. ESTIMATION PROCEDURES and VI. RELIABILITY STATEMENT)

Summary expenditure variables in this file can be combined to derive weekly estimates for broad consumption categories. Demographic characteristics, such as family size, refer to the CU status on the date of the interview. Income variables contain annual values, covering the 12 months prior to the date of the interview. 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. SUMMARY EXPENDITURE DATA

The variables FOODTOT through HOUSKEEP contain summary expenditure data. They are all BLS derived. The UCCs comprising each summary expenditure variable are listed below the variable description. UCCs may not be represented in all Diary quarters. When UCCs are added or deleted to the summary variable definition, the quarter in which the addition (deletion) to the summary expenditure variable occurs is denoted by a leading superscript directly after the UCC code in the “Changes to the YYYY Microdata section”. For example, N101<UCC> or D101<UCC> identifies a new or deleted UCC for a given summary expenditure variable beginning in Q101.

2. ***MEMBER CHARACTERISTICS AND INCOME (MEMB) FILE***

The "MEMB" file, also referred to as the "Member Characteristics and Income" file, contains selected characteristics for each CU member, including identification of relationship to reference person. Characteristics for the reference person and spouse appear on both the MEMB file and FMLY file. Demographic characteristic data, such as age of CU member, refer to the member status at the placement of each diary. Income data are collected for all CU members over 13 years of age. Income taxes withheld and pension and retirement contributions are shown both annually and as deductions from the member's last paycheck. Income variables contain annual values for the 12 months prior to the interview month. 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.

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

In the "EXPN" file, each expenditure recorded by a CU in a weekly diary is identified by UCC, 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 no expenditure was recorded for the item(s) represented by a UCC, then there is no record for the UCC on file.

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 and contains the same annual and point-of-placement data. It was created to facilitate computer processing when linking CU income and demographic characteristic data with EXPN expenditure data. As such, the file structure is similar to EXPN. Each characteristic and income item is identified by UCC (See Section XIII.B for a listing of UCCs). 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.

5. ***IMPUTED INCOME FILE (DTID)***

As a result of the introduction of multiply imputed income data in the Consumer Expenditure Survey, the Imputed DTAB (DTID) file is now on the Microdata. It is very similar to the DTAB file, except that the variable “IMPNUM” will indicate the number (1-5) of the imputation variant of the income variable and it only contains UCCs from variables subject to income imputation.

6. ***PROCESSING FILES***

a. Dstub file

X:\Programs\Dstub2010.txt

The Dstub file shows the aggregation scheme used in the published consumer expenditure tables. It is formatted as follows:

DESCRIPTION	START POSITION	FORMAT
Type: represents whether information in this line contains aggregation data or not	1	CHAR(1)
Level: aggregation level (lowest number is highest level of aggregation)	4	CHAR(1)
Title: title of the line item	7	CHAR(60)
UCC: UCC number in the EXPN or DTAB file	70	CHAR(6)
Survey: Indicates survey source (D = Diary, G = Aggregated item)	80	CHAR(1)
Group: Indicates if the item is an expenditure, income, or asset	86	CHAR(7)

Note: this file is an internal BLS file used for processing expenditures. It has other information that may be ignored by users of the public use data.

b. UCC file

X:\DIARY10\UCCD10.TXT

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

DESCRIPTION	START POSITION	FORMAT
UCC	1	CHAR(6)
UCC title (See Section XIII.A. EXPENDITURE UCCS ON EXPN FILE and XIII.B. INCOME AND RELATED UCCS ON DTAB FILE for a list of UCCs and their full titles by file—expenditure (EXPN) or income (DTAB).)	8	CHAR(50)

c. Sample program files

X:\Programs\SAS\Diary Mean and SE.sas
X:\Programs\SAS\Integrated Mean and SE.sas
X:\Programs\STATA\Diary Table2010.do

The Diary Mean and SE program file contains the computer program used in Section VII.A. SAMPLE PROGRAM of the documentation. This file has been created to provide programming assistance.

In addition to the Diary Mean and SE.sas program, there are additional sample programs in the Programs folder to provide assistance using different files. The Integrated Mean and SE.sas program is used using data from the Diary and Interview files to create the means and standard errors using the same methods as the published tables. The Diary Table do-file is a STATA program that produces the same Diary Mean estimates.

Note: Estimates from the programs will not match the published tables exactly due to topcoding in the public-use data.

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 calculated 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 critical value and topcode amount. If an observation is topcoded, the flag variable assigned to that observation is set to 'T'.

Since the critical value and 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 and be applied at a different starting point. By topcoding values in this manner, 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.

A. CU CHARACTERISTICS AND INCOME FILE (FMLY)

The following table lists FMLY file variables that are subject to topcoding as well as their associated critical values and topcode values. For multiply imputed income variables, it is possible for an upper topcode value to be less than the upper critical value or for a lower topcode value to be greater than the lower critical value.

Variable	Description	2010 Upper Critical Value	2010 Lower Critical Value	2010 Upper Topcode Value	2010 Lower Topcode Value
ADDFEDX	Amount of Federal income tax paid in addition to that withheld	24,000	NA	52,023	NA
ADDOTHX	Amount of other taxes paid but not reported elsewhere	10,000	NA	16,143	NA
ADDSTAX	Amount of state and local income tax paid in addition to that withheld	6,000	NA	21,770	NA
ALIOTHX	Amount received from regular contributions by all CU members	30,000	NA	83,000	NA
ALIOTHXM	Amount received from regular contributions by all CU members	30,000	NA	46,150	NA
CHDLMPX	Amount received by all CU members for a lump sum child support payment in last 12 months	4,800	NA	15,240	NA

Variable	Description	2010 Upper Critical Value	2010 Lower Critical Value	2010 Upper Topcode Value	2010 Lower Topcode Value
CHDOTHX	Amount received by all CU members in last 12 months for other child support	15,600	NA	21,510	NA
CHDOTHXM	Amount received by all CU members in last 12 months for other child support	15,600	NA	17,598	NA
DIVX	Amount received from dividends, royalties, estates, or trusts	75,000	NA	139,400	NA
DIVXM	Amount received from dividends, royalties, estates, or trusts	75,000	NA	64,696	NA
FEDREFX	Amount of refund from Federal income tax	9,902	NA	15,046	NA
INSREFX	Amount of refund from insurance policies	4,800	NA	18,800	NA
INTX	Amount received from interest on savings accounts, or bonds	35,000	NA	149,307	NA
INTXM	Amount received from interest on savings accounts, or bonds	35,000	NA	46,372	NA
LUMPX	Amount from lump sum payments from estates, trusts, royalties, alimony, child support, prizes, games of chance, or persons outside CU	124,000	NA	321,333	NA
OCCEXPNX	Amount paid by CU for occupational expenses, last 12 months	5,000	NA	15,140	NA
OTHINX	Amount from other money income, including money from care of foster children, cash scholarships and fellowships, or stipends, not based on working	32,000	NA	88,667	NA
OTHINXM	Amount from other money income, including money from care of foster children, cash scholarships and fellowships, or stipends, not based on working	32,000	NA	48,069	NA
OTHREFX	Amount of refund from other sources, including any other taxes	1,000	NA	3,327	NA
PENSIONM	Amount received from pensions or annuities from private companies, military or government, IRA or Keogh	72,000	NA	76,637	NA
PENSIONX	Amount received from pensions or annuities from private companies, military or government, IRA or Keogh	72,000	NA	146,081	NA
PTAXREFX	Amount of refund from property taxes	1,800	NA	5,300	NA
ROOMX	Amount of net income or loss received from roomers or boarders	30,600	NA	73,200	NA

Variable	Description	2010 Upper Critical Value	2010 Lower Critical Value	2010 Upper Topcode Value	2010 Lower Topcode Value
ROOMXM	Amount of net income or loss received from roomers or boarders	30,600	NA	41,182	NA
SALEX	Amount received from sale of household furnishings, equipment, clothing, jewelry, pets or other belongings, excluding sale of vehicles or property	7,000	NA	27,400	NA
SSREFX	Amount of refund from overpayment on Social Security	2,040	NA	5,725	NA
STATREFX	Amount of refund from state or local income tax	2,500	NA	4,639	NA
TAXPROPX	Amount of personal property taxes paid but not reported elsewhere	1,200	NA	2,258	NA

Some income variables that are subject to topcoding are constructed by summing up the values of “lower level” MEMB or FMLY file component variables. These variables are not topcoded by the conventional method of replacement with a topcode value. Instead the variables' components are summed normally and the variables are flagged as topcoded if one of their component variables is topcoded.

Following are the income variables that are calculated using values of their component variables. (See the descriptions of each variable in the diary data dictionary for a list of component variables.)

EARNX	Amount of CU income from earnings before taxes
FBSNSXM, FBSNSX1-5 FBSNSX	Amount of income from non-farm business
FFARMXM, FFARMX1-5 FFARMX	Amount of income or loss received from own farm
FFEDTXX	Amount of Federal tax deducted from last pay, annualized for all CU members
FGVXM, FGVX1-5 FGVX	Amount of government retirement deducted from last pay, annualized for all CU members
FINCAFTM, FINCAFT1-5 FINCAFTX	Amount of CU income after taxes
FINCBEFM, FINCBEF1-5 FINCBEFX	Amount of CU income before taxes
FIRAX	Amount of money placed in individual retirement plan
FJSSDEDM, FJSSDED1-5 FJSSDEDX	Estimated amount of annual Social Security contribution
FPVTXM FPVTX	Amount of private pension fund deducted from last pay, annualized for all CU members
FRRXM FRRX	Amount of Railroad Retirement deducted from last pay, annualized for all CU members
FSTATXXM,	Amount of State and local income taxes deducted from last pay, annualized for all CU

FSTATXX1-5 FSTATXX	members
FWAGEXM, FWAGEX1-5 FWAGEX	Amount received from wage and salary income before deduction
OTHRECX	Amount of other money receipts excluded from family income
PERSTAXM, PERSTAX1-5 PERSTAX	Amount of personal taxes paid

Here are some examples of situations that may occur. The value for the variable FBSNSX (family income from nonfarm business) is computed as the sum of the values reported for the variable BSNSX (member income from nonfarm business) from the MEMB file. BSNSX is subject to topcoding beyond the critical value of $200,000 (-$9,999). The topcode value for BSNSX is $298,000 (-$115,703).

		BSNSX		**FBSNSX**	
CU		*REPORTED*	*AFTER TOPCODING*	*VALUE*	*FLAGGED AS TOPCODED?*
CU 1:	MEMB1	$145,000	$145,000		
	MEMB2	145,000	145,000		
	MEMB3	20,000	20,000	310,000	No
CU 2:	MEMB1	354,000	298,000		
	MEMB2	-15,000	-115,703		
	MEMB3	-29,000	-115,703	66,594	Yes
CU 3	MEMB1	205,000	298,000		
	MEMB2	130,000	130,000	428,000	Yes
CU 4	MEMB1	140,000	140,000		
	MEMB2	140,000	140,000		
	MEMB3	-300,000	-115,703	164,297	Yes

While CUs 1 and 2 each originally report a total of $310,000 for all members in BSNSX, topcoding is done only on the values reported by the members 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 ($428,000) is higher than the amount that is reported ($335,000). The case of CU4 demonstrates that the reported value for FBSNSXM can be negative, while the topcoded value can be positive. The reverse can also occur.

The value of the variable, STATE, which identifies state of residence, must be suppressed for some observations to meet the Census Disclosure Review Board's criterion that the smallest geographically identifiable area have a population of at least 100,000. STATE data were evaluated vis-à-vis variables POPSIZE, REGION, and BLS_URBN, which show the population size of the geographic area that is sampled, the four Census regions, and the urban/rural status respectively. Some STATE codes were suppressed because, in combination with these

variables, they could be used to identify areas of 100,000 or less. On approximately 13 percent of the records on the FMLY files the STATE variable is blank.

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.

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

* indicates that the STATE code has been suppressed for all sampled CUs in that state.

** indicates that the STATE code has been suppressed for some sampled CUs in that state.

R indicates that either all observations from this state have been re-coded or all strata[1] 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 stratum[1] of observations from this state includes "re-codes" from other states.

R* indicates that the STATE code has been suppressed for some sampled CUs in that state and, either STATE has been re-coded or the state includes "re-codes" from other states in all strata[1].

RR** indicates that the STATE code has been suppressed for some sampled CUs in that state and, either STATE has been re-coded or the state includes "re-codes" from other states in at least one stratum[1].

[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 table lists MEMB file variables that are subject to topcoding as well as their associated critical values and topcode values. For multiply imputed income variables, it is

possible for an upper topcode value to be less than the upper critical value or for a lower topcode value to be greater than the lower critical value.

Variable	Description	2010 Upper Critical Value	2010 Lower Critical Value	2010 Upper Topcode Value	2010 Lower Topcode Value
AGE	Age of member	82	NA	87	NA
ANFEDTXM	Annual amount of Federal income tax deducted from pay	25,526	NA	45,292	NA
ANFEDTXX	Annual amount of Federal income tax deducted from pay	25,526	NA	44,639	NA
ANGVX	Annual amount of government retirement deducted from pay	9,655	NA	12,741	NA
ANGVXM	Annual amount of government retirement deducted from pay	9,655	NA	12,741	NA
ANPVTX	Annual amount of private pension fund deducted from pay	19,960	NA	29,445	NA
ANPVTXM	Annual amount of private pension fund deducted from pay	19,960	NA	30,081	NA
ANSTATXM	Annual amount of state and local income taxes deducted from pay	9,100	NA	18,044	NA
ANSTATXX	Annual amount of state and local income taxes deducted from pay	9,100	NA	17,757	NA
BSNSX	Amount of income or loss received from nonfarm business	200,000	-9,999	298,000	-115,703
BSNSXM	Amount of income or loss received from nonfarm business	200,000	-9,999	122,201	-21,852
FARMX	Amount of income or loss received from own farm	NA	-9,999	NA	-28,055
FARMXM	Amount of income or loss received from own farm	NA	-9,999	NA	-16,029
FEDTXX	Amount of Federal income tax deducted from last pay	1,200	NA	3,405	NA
GROSPAYX	Amount of last gross pay	6,800	NA	14,871	NA
GVX	Amount of government retirement deducted from last pay	850	NA	2,468	NA
IRAX	Amount of money placed in an individual retirement plan	25,000	NA	65,333	NA
JSSDEDX	Estimated annual Social Security contribution	8,797	NA	14,714	NA
JSSDEDXM	Estimated annual Social Security contribution	8,797	NA	9,648	NA
PVTX	Amount of private pension fund deducted from last pay	1,200	NA	6,696	NA
SLFEMPSM	Amount of self-employment Social Security contributions	17,593	NA	9,766	NA
SLFEMPSS	Amount of self-employment Social Security contributions	17,593	NA	20,733	NA
STATXX	Amount of state and local income taxes deducted from last pay	430	NA	1,242	NA

Variable	Description	2010 Upper Critical Value	2010 Lower Critical Value	2010 Upper Topcode Value	2010 Lower Topcode Value
WAGEX	Amount received from wage and salary income before deductions	150,000	NA	448,676	NA
WAGEXM	Amount received from wage and salary income before deductions	150,000	NA	261,983	NA

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 WAGEXM (annual wage and salary income) to derive ANFEDTXM, ANGVXM, ANPVTXM, ANRRXM, and ANSTATXM, 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) ANFEDTXM = (WAGEXM (FEDTXX/GROSPAYX)).

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

(2) WAGEXM = (ANFEDTXM (GROSPAYX/FEDTXX)).

In the above example, a problem with disclosure may arise when neither ANFEDTXM, GROSPAYX, nor FEDTXX (calculation components) are topcoded, *but WAGEXM is*. In this situation WAGEXM can be recalculated to obtain its original value by inserting the non-topcoded values into equation (2) and solving it. In order to prevent this, the non-topcoded terms in equation (2) will be suppressed (blanked out) and their associated flags will be assigned a value of 'T'.

The following chart describes in detail the specific rules that are applied to prevent the potential disclosure outlined above.

If WAGEXM is greater than the critical value but ANFEDTXM, GROSPAYX, and FEDTXX are not, then the values for ANFEDTXM, GROSPAYX, and FEDTXX are suppressed and their flag variables are assigned a value of 'T'.

If WAGEXM is greater than the critical value but ANGVXM, GROSPAYX, and GVX are not, then the values for ANGVXM, GROSPAYX, and GVX are suppressed and their flag variables assigned a value of 'T'.

If WAGEXM is greater than the critical value but ANPVTXM, GROSPAYX, and PVTX are not, then the values for ANPVTXM, GROSPAYX, and PVTX are suppressed and their flag variables assigned a value of 'T'.

If WAGEXM is greater than the critical value but ANRRXM, GROSPAYX, and RRX are not, then the values for ANRRXM, GROSPAYX, and RRX are suppressed and their flag variables assigned a value of 'T'.

If WAGEXM is greater than the critical value but ANSTATXM, GROSPAYX, and STATXX are not, then the values for ANSTATXM, GROSPAYX, and STATXX are suppressed and their flag variables assigned a value of 'T'.

The same special suppression for MEMB file variables occurs with the original (pre-income imputation) variables that correspond to the variables noted above (WAGEX, ANFEDTXX, etc)

C. DETAILED EXPENDITURE FILE (EXPN)

The following table lists UCCs for which the EXPN variable COST is subject to topcoding as well as their associated critical values and topcode values (rounded to the nearest dollar). If the value of COST is greater (less) than the designated critical values for the above UCCs, COST is set to the topcode value and the associated flag variable, COST_, is set to 'T'.

Variable	Description	2010 Upper Critical Value	2010 Lower Critical Value	2010 Upper Topcode Value	2010 Lower Topcode Value
001000	Purchase price of stocks, bonds, mutual funds	1,693	NA	4,169	NA
009000	Mortgage payment including coop	3,360	NA	4,492	NA
210110	Rent of dwelling, includes parking fees	2,095	NA	3,242	NA
210210	Lodging away from home	869	NA	1,897	NA
210310	Housing for someone at school	0	NA	208	NA
220400	Purchase of property	0	NA	3,929	NA
550320	Medical equipment for general use	65	NA	142	NA
550330	Supportive convalescent or medical equipment	110	NA	164	NA
560110	Physicians' services	223	NA	460	NA
560210	Dental services	950	NA	1,450	NA
560310	Eyecare services	650	NA	1,663	NA
560330	Lab tests and x-rays	195	NA	317	NA
560400	Service by professionals other than physicians	260	NA	508	NA
570000	Hospital care not specified	700	NA	1,219	NA
570220	Nursing or convalescent home care	4,745	NA	6,386	NA
570230	Other medical care service	150	NA	497	NA
570901	Rental of medical equipment	26	NA	40	NA

D. INCOME FILE (DTAB)

The following table lists UCCs for which the DTAB variable AMOUNT is subject to topcoding as well as their associated critical values and topcode values (rounded to the nearest dollar). If the value of AMOUNT is greater (less) than the designated critical values for the above UCCs, AMOUNT is set to the topcode value and the associated flag variable, AMOUNT_, is set to 'T'

Variable	Description	2010 Upper Critical Value	2010 Lower Critical Value	2010 Upper Topcode Value	2010 Lower Topcode Value
900040	Amount received from pensions or annuities	72,000	NA	76,637	NA

Variable	Description	2010 Upper Critical Value	2010 Lower Critical Value	2010 Upper Topcode Value	2010 Lower Topcode Value
900050	Amount received from regular income from dividends, royalties, estates or trusts	75,000	NA	64,696	NA
900060	Amount received from net income or loss received from roomers or boarders	30,600	NA	41,182	NA
900080	Amount received from interest on savings accounts or bonds	35,000	NA	46,372	NA
900131	Amount received from other child support payments	15,600	NA	17,598	NA
900132	Amount received from other regular contributions, including alimony	30,000	NA	46,150	NA
900140	Amount received from other money income	32,000	NA	48,069	NA
910000	Amount received from lump sum payments from estates, trusts, etc.	124,000	NA	321,333	NA
910010	Amount received from money from sale household furnishings etc.	7,000	NA	27,400	NA
910020	Amount of overpayment on Social Security	2,040	NA	5,725	NA
910030	Amount of refund from insurance policies	4,800	NA	18,800	NA
910040	Amount of refunds from property taxes	1,800	NA	5,300	NA
910041	Amount received from lump sum child	4,800	NA	15,240	NA
950001	Amount received from federal income tax refunds	NA	-9,902	NA	-15,046
950003	Amount of additional federal income tax paid (not deducted)	24,000	NA	52,023	NA
950011	Amount received from state/local income tax refunds	NA	-2,500	NA	-4,639
950013	Amount of additional state/local income tax paid (not deducted)	6,000	NA	21,770	NA
950021	Amount of other taxes paid	10,000	NA	16,143	NA
950022	Amount of personal property taxes paid	1,200	NA	2,258	NA
950023	Amount of other tax refund received from other sources	NA	-1,000	NA	-3,327

[1] ADDFEDX (amount of Federal tax paid in addition to that withheld) and FFEDTXX (Federal tax withheld from last pay annualized for all CU members) are mapped to UCCs 950003 and 950002, respectively, as separate records. Records for UCC 950002 that represent FFEDTXX are topcoded through their components (ANFEDTXM) at the MEMB level and thus, these records will not have a DTAB critical value. DTAB records for UCC 950003 that represent ADDFEDX are topcoded for all amounts greater than $24,000.

[2] ADDSTAX (amount of state and local taxes paid in addition to that withheld) and FSTATXX (state and local income tax deduction from last pay annualized for all CU members) are mapped to UCCs 950013 and 950012, respectively, as separate records. Records for UCC 950012 that represent FSTATXX are topcoded through their components (ANSTATXM) at the MEMB level and thus, these records will not have a DTAB critical value. Create the DTAB VALUE

field for these records by dividing FSTATXX by 12. If FSLTAXX is topcoded, then set VALUE_ to 'T'. DTAB records for UCC 950013 that represent ADDSTAX are topcoded for all amounts greater than $6,000.

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. CU CHARACTERISTICS AND INCOME FILE on topcoding of FMLY variables.)

UCC	FMLY variable	Description
800910	FGVXM, FGVX	Amount of government retirement deducted from last pay, annualized for all CU members
800920	FRRXM ,FRRX	Amount of Railroad Retirement deducted from last pay, annualized for all CU members
800931	FPVTXM, FPVTX	Amount of private pension fund deducted from last pay, annualized for all CU members
800932	FIRAX	Amount of money placed in individual retirement plan
800940	FJSSDEDM, FJSSDED1-5, FJSSDEDX	Estimated amount of annual Social Security contribution
900000	FWAGEXM, FWAGEX1-5, FWAGEX	Amount received from wage and salary income before deduction
900010	FBSNSXM, FBSNSX1-5, FBSNSX	Amount of income from non-farm business
900020	FFARMXM, FFARMX1-5, FFARMX	Amount of income or loss received from own farm
980000	FINCBEFM, FINCBEF1-5, FINCBEFX	Amount of CU income before taxes
980070	FINCAFTM, FINCAFT1-5, FINCAFTX	Amount of CU income after taxes

V. ESTIMATION PROCEDURE

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 2010* used an integration methodology which incorporated information from *both* Diary and Interview Surveys. Diary data users will not be able to match published CE estimates because of this. 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

$\overline{X}_{(S,k)(q,r)}$ 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 $X_{(S,k)(12,6)}$.

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

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

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

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, $\overline{X}_{(S,k)(12,6)}$, 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

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

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 $\overline{X}_{(S,k)(q,k)}$, (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:

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

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 $\overline{X}_{(S,k)(q,r)}$.

The estimates for total ($X_{(S,k)(q,r)}$) and mean expenditures ($\overline{X}_{(S,k)(q,r)}$) 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 $\overline{X}_{(S,k)(q,r)}$ 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 $\overline{Z}_{(S,r)}$ 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

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

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

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.

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:

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

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

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

and

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

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

2. ***STANDARD ERROR OF THE MEAN***

The standard error of the mean, *S.E.*($\overline{X}$), is defined as the square root of the variance of the mean. *S.E.*($\overline{X}$), is used to obtain confidence intervals that evaluate how close the estimate may be to the true population mean. A 95 percent confidence interval can be constructed around an estimate, bounded by values 1.96 times the standard error less than and greater than the estimate. For example, the average weekly expenditure for beef for All CUs in 2010 was $4.17. The standard error for this estimate is $0.19. Hence, the 95 percent confidence interval around this estimate is from $3.80 to $4.54. Therefore, we could conclude with 95 percent confidence that the mean weekly expenditures for beef all CUs in 2010 lies within the interval $3.80 to $4.54.

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 2010 the estimated average weekly expenditures for total food for CUs in the *$30,000 to $39,999* income range is $88.96 and the estimate for CUs in the *$40,000 to $49,999* income range is $100.55. The apparent difference between the two mean expenditures is $100.55 – $88.96 = $11.59. The standard error on the estimate of $88.96 is $2.96 and the estimated standard error for the $100.55 estimate is $3.41. The standard error (S.E.) of a difference is approximately equal to

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

where

$$V(\overline{X}_i) = \left(S.E.(\overline{X}_i)\right)^2$$

This assumes that $\overline{x}_1$ and $\overline{x}_2$ are disjoint subsets of the population. Hence, the standard error of the difference in food expenditures between CUs in the *$30,000 to $39,999* and in the *$40,000 to $49,999* income ranges is about

$$\sqrt{\left((2.96)^2 + (3.41)^2\right)} = 4.52$$

This means that the 95 percent confidence interval around the difference is from $2.73 to $20.45. Since this interval does not include zero, we can conclude with 95 percent confidence that the mean weekly food expenditures for the *$40,000 to $49,999* income group is greater than the mean weekly food expenditures for the *$30,000 to $39,999* income 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

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

where

$$V(\overline{X}_i) = \left(S.E.(\overline{X}_i)\right)^2$$

and where *r* is the correlation coefficient between $\overline{x}_1$ and $\overline{x}_2$. The correlation coefficient is generally no greater than 0.2 for CE estimates.

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 shows usage of the SAS datasets available on the SAS CD-ROM. A program written in SAS but utilizing the ASCII datasets is present on the ASCII CD-ROM but will not be referenced here. Refer to the output file on the CD 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 the diary data dictionary.

This program produces a table of selected expenditures by income class of the Consumer Unit (CU). The first section reads in the processing file and manipulates it into a usable form suitable for formatting an expenditure table. The second section of the program extracts the relevant variables from the FMLY files, while the third section extracts the expenditure and income data from the EXPN and DTAB files. These three datasets are then used along with the Dstub processing file 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.

```
380    /*************************************************************************/
381    /* PROGRAM NAME:  CEX DIARY SURVEY SAMPLE PROGRAM (SAS)                  */
382    /* LOCATION: D:\PROGRAMS                                                 */
383    /* FUNCTION: CREATE A DIARY SURVEY EXPENDITURE TABLE BY INCOME CLASS USING */
384    /*           MICRODATA FROM THE BUREAU OF LABOR STATISTIC'S CONSUMER      */
385    /*           EXPENDITURE SURVEY.                                         */
386    /*                                                                       */
387    /* WRITTEN BY:  ERIC KEIL                                                */
388    /* MODIFICATIONS:                                                        */
389    /* DATE-      MODIFIED BY-      REASON-                                  */
390    /* -----      ------------      -------                                  */
391    /* 03/21/02   ERIC KEIL         IMPROVE EFFICIENCY                       */
392    /* 10/22/03   ERIC KEIL         UPDATE FOR 2002 DATA                     */
393    /* 11/20/03   ERIC KEIL         INCLUDE ROUTINE TO AGGREGATE EASIER      */
394    /*                                                                       */
395    /*                                                                       */
396    /*  FOR SAS VERSION 8 OR HIGHER                                          */
397    /*                                                                       */
398    /*  DATA AND INPUT FILES USED IN THIS SAMPLE PROGRAM WERE UNZIPPED        */
399    /*  OR COPIED TO THE LOCATIONS BELOW:                                    */
400    /*                                                                       */
401    /*  DIARY DATA -- C:\2009_CEX\DIARY09                                    */
402    /*  DSTUB2008.TXT -- C:\2009_CEX\Programs                                */
403    /*                                                                       */
404    /*************************************************************************/
405
406    /*Enter Data Year*/
407      %LET YEAR = 2010;
408    /*Enter location of the unzipped microdata file*/
409      %LET DRIVE = C:\2010_CEX;
410
411    /*************************************************************************/
412    /* STEP1: READ IN THE STUB PARAMETER FILE AND CREATE FORMATS             */
413    /* --------------------------------------------------------------------- */
414    /* 1 CONVERTS THE STUB PARAMETER FILE INTO A LABEL FILE FOR OUTPUT       */
415    /* 2 CONVERTS THE STUB PARAMETER FILE INTO AN EXPENDITURE AGGREGATION FILE */
416    /* 3 CREATES FORMATS FOR USE IN OTHER PROCEDURES                         */
417    /*************************************************************************/
418
419
420  %LET YR1 = %SUBSTR(&YEAR,3,2);
421  LIBNAME D&YR1 "&DRIVE\DIARY&YR1";
NOTE: Libref D10 was successfully assigned as follows:
      Engine:        V9
      Physical Name: C:\2010_CEX\DIARY10
422
423
424  DATA STUBFILE (KEEP= COUNT TYPE LEVEL TITLE UCC SURVEY GROUP LINE);
425    INFILE "&DRIVE\PROGRAMS\DSTUB&YEAR..TXT"
426    PAD MISSOVER;
427    INPUT @1 TYPE $1. @ 4 LEVEL $1. @7 TITLE $CHAR60. @70 UCC $6.
428          @80 SURVEY $1. @86 GROUP $7.;
429    IF (TYPE = '1');
430    IF GROUP IN ('CUCHARS' 'FOOD' 'EXPEND' 'INCOME');
431    IF SURVEY = 'T' THEN DELETE;
432      RETAIN COUNT 9999;
433      COUNT + 1;
434      LINE = PUT(COUNT, $5.)||LEVEL ;
WARNING: Variable COUNT has already been defined as numeric.
435      /* READS IN THE STUB PARAMETER FILE AND CREATES LINE NUMBERS FOR UCCS */
436      /* A UNIQUE LINE NUMBER IS ASSIGNED TO EACH EXPENDITURE LINE ITEM     */
437  RUN;

NOTE: The infile "C:\2010_CEX\PROGRAMS\DSTUB2010.TXT" is:
      File Name=C:\2010_CEX\PROGRAMS\DSTUB2010.TXT,
      RECFM=V,LRECL=256

NOTE: 794 records were read from the infile "C:\2010_CEX\PROGRAMS\DSTUB2010.TXT".
      The minimum record length was 91.
      The maximum record length was 92.
```

Sets the calendar year and drive used as macro variables that can be used throughout the program.

Reads in the aggregation stub file and dynamically creates numbers associated with each expenditure line item.

Note: This aggregation file can be modified to accommodate any customized aggregation scheme.

One needs only to make sure that the column start positions in the file match the start positions in the input statement.

```
NOTE: The data set WORK.STUBFILE has 480 observations and 8 variables.
NOTE: DATA statement used (Total process time):
      real time           0.04 seconds
      cpu time            0.01 seconds


438
439
440  DATA AGGFMT1 (KEEP= UCC LINE LINE1-LINE10);
441    SET STUBFILE;
442    LENGTH LINE1-LINE10 $6.;
443      ARRAY LINES(9) LINE1-LINE9;
444        IF (UCC > 'A') THEN
445          LINES(SUBSTR(LINE,6,1)) = LINE;
446        RETAIN LINE1-LINE9;
447        IF (UCC < 'A')  THEN
448          LINE10 = LINE;
449    IF (LINE10);
450  RUN;

NOTE: Character values have been converted to numeric values at the places given by:
(Line):(Column).
      445:15   449:7
NOTE: There were 480 observations read from the data set WORK.STUBFILE.
NOTE: The data set WORK.AGGFMT1 has 354 observations and 12 variables.
NOTE: DATA statement used (Total process time):
      real time           0.04 seconds
      cpu time            0.01 seconds


451
452
453  PROC SORT DATA= AGGFMT1 (RENAME=(LINE= COMPARE));
454    BY UCC;
455      /* MAPS LINE NUMBERS TO UCCS */
456  RUN;

NOTE: There were 354 observations read from the data set WORK.AGGFMT1.
NOTE: The data set WORK.AGGFMT1 has 354 observations and 12 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.06 seconds
      cpu time            0.03 seconds


457
458
459  PROC TRANSPOSE DATA= AGGFMT1 OUT= AGGFMT2 (RENAME=(COL1= LINE));
460    BY UCC COMPARE;
461    VAR LINE1-LINE10;
462  RUN;

NOTE: There were 354 observations read from the data set WORK.AGGFMT1.
NOTE: The data set WORK.AGGFMT2 has 3540 observations and 4 variables.
NOTE: PROCEDURE TRANSPOSE used (Total process time):
      real time           0.04 seconds
      cpu time            0.01 seconds


463
464
465  DATA AGGFMT (KEEP= UCC LINE);
466    SET AGGFMT2;
467      IF LINE;
468      IF SUBSTR(COMPARE,6,1) > SUBSTR(LINE,6,1) OR COMPARE=LINE;
469      /* AGGREGATION FILE. EXTRANEOUS MAPPINGS ARE DELETED            */
470      /* PROC SQL WILL AGGANGE LINE#/UCC PAIRS FOR USE IN PROC FORMAT */
471  RUN;

NOTE: Character values have been converted to numeric values at the places given by:
(Line):(Column).
      467:8
```

Subsequent program steps manipulate the aggregation stub file into a dataset that associates UCCs with line numbers.

```
NOTE: There were 3540 observations read from the data set WORK.AGGFMT2.
NOTE: The data set WORK.AGGFMT has 1420 observations and 2 variables.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


472
473
474  PROC SQL NOPRINT;
475    SELECT UCC, LINE, COUNT(*)
476    INTO  :UCCS SEPARATED BY " ",
477          :LINES SEPARATED BY " ",
478          :CNT
479    FROM AGGFMT;
NOTE: The query requires remerging summary statistics back with the original data.
480    QUIT;
NOTE: PROCEDURE SQL used (Total process time):
      real time           0.04 seconds
      cpu time            0.01 seconds


481  RUN;
482
483
484  %MACRO MAPPING;
485    %DO  I = 1  %TO  &CNT;
486      "%SCAN(&UCCS,&I,%STR( ))" = "%SCAN(&LINES,&I,%STR( ))"
487    %END;
488  %MEND MAPPING;
489
490
491  DATA LBLFMT (RENAME=(LINE= START TITLE= LABEL));
492    SET STUBFILE (KEEP= LINE TITLE);
493    RETAIN FMTNAME 'LBLFMT' TYPE 'C';
494    /* LABEL FILE. LINE NUMBERS ARE ASSIGNED A TEXT LABEL */
495    /* DATASET CONSTRUCTED TO BE READ INTO A PROC FORMAT  */
496  RUN;

NOTE: There were 480 observations read from the data set WORK.STUBFILE.
NOTE: The data set WORK.LBLFMT has 480 observations and 4 variables.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.00 seconds


497
498
499  PROC FORMAT;
500
501    VALUE $AGGFMT (MULTILABEL)
502      %MAPPING
503      OTHER= 'OTHER';
NOTE: Format $AGGFMT is already on the library.
NOTE: Format $AGGFMT has been output.
504      /* CREATE AGGREGATION FORMAT */
505
506
507    VALUE $INC (MULTILABEL)
508      '01' = '01'
509      '01' = '10'
510      '02' = '02'
511      '02' = '10'
512      '03' = '03'
513      '03' = '10'
514      '04' = '04'
515      '04' = '10'
516      '05' = '05'
517      '05' = '10'
518      '06' = '06'
519      '06' = '10'
```

Creates a Dataset that can be used to associate titles with line numbers with a format procedure.

Formats:

Puts the aggregation scheme into a SAS format.

Puts the income groupings into a SAS format.

Note: The multilabel option is necessary in the aggregation format and income format since multiple mappings occur. This option is available in SAS V8 or higher.

```
520      '07' = '07'
521      '07' = '10'
522      '08' = '08'
523      '08' = '10'
524      '09' = '09'
525      '09' = '10';
NOTE: Format $INC is already on the library.
NOTE: Format $INC has been output.
526      /* CREATE INCOME CLASS FORMAT */
527  RUN;

NOTE: PROCEDURE FORMAT used (Total process time):
      real time           3.21 seconds
      cpu time            3.20 seconds


528
529
530  PROC FORMAT LIBRARY= WORK  CNTLIN= LBLFMT;
NOTE: Format $LBLFMT has been output.
531    /* CREATE LABEL FILE FORMATS */
532  RUN;

NOTE: PROCEDURE FORMAT used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds

NOTE: There were 480 observations read from the data set WORK.LBLFMT.

533
534
535    /*************************************************************************/
536    /* STEP2: READ IN ALL NEEDED DATA FROM THE CD-ROM                        */
537    /* --------------------------------------------------------------------- */
538    /* 1 READ IN THE DIARY FMLY FILES                                        */
539    /* 2 READ IN THE DIARY EXPM AND DTAB FILES                               */
540    /* 3 MERGE FMLY AND EXPENDITURE FILES TO DERIVE WEIGHTED EXPENDITURES    */
541    /*************************************************************************/
542
543
544  DATA FMLY (KEEP = NEWID INCLASS REPWT1-REPWT45);
545    SET D&YR1..FMLD&YR1.1
546        D&YR1..FMLD&YR1.2
547        D&YR1..FMLD&YR1.3
548        D&YR1..FMLD&YR1.4;

550        /* READ IN FMLY FILE DATA */
551
552      ARRAY REPS_A(45) WTREP01-WTREP44 FINLWT21;
553      ARRAY REPS_B(45) REPWT1-REPWT45;
554
555        DO i = 1 TO 45;
556        IF REPS_A(i) > 0 THEN
557           REPS_B(i) = (REPS_A(i) / 4);
558           ELSE REPS_B(i) = 0;
559        END;
560        /* ADJUST WEIGHTS TO COMPENSATE FOR HAVING FOUR QUARTERS OF DATA */
561  RUN;

NOTE: There were 3725 observations read from the data set D10.FMLD101.
NOTE: There were 3619 observations read from the data set D10.FMLD102.
NOTE: There were 3511 observations read from the data set D10.FMLD103.
NOTE: There were 3441 observations read from the data set D10.FMLD104.
NOTE: The data set WORK.FMLY has 14296 observations and 47 variables.
NOTE: DATA statement used (Total process time):
      real time           0.35 seconds
      cpu time            0.31 seconds


562
563
```

Puts the titles into a SAS format for use in the final output.

Reads in the necessary variables from the fmly files. Newid is the code given to a consumer unit each time it participates. Finlwt21 and Wtrep01-Wtrep44 are weight variables 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.

Lines 555-559 adjust the weights so that they will sum up to US populations.

```
564
565  DATA EXPEND (KEEP = NEWID UCC COST);
566    SET D&YR1..DTBD&YR1.1 (RENAME=(AMOUNT=COST))
567        D&YR1..DTBD&YR1.2 (RENAME=(AMOUNT=COST))
568        D&YR1..DTBD&YR1.3 (RENAME=(AMOUNT=COST))
569        D&YR1..DTBD&YR1.4 (RENAME=(AMOUNT=COST))
570        D&YR1..EXPD&YR1.1
571        D&YR1..EXPD&YR1.2
572        D&YR1..EXPD&YR1.3
573        D&YR1..EXPD&YR1.4;
574    BY NEWID;
575    /* READ IN INCOME AND EXPENDITURE DATA */
576  RUN;

NOTE: There were 63052 observations read from the data set D10.DTBD101.
NOTE: There were 61362 observations read from the data set D10.DTBD102.
NOTE: There were 59099 observations read from the data set D10.DTBD103.
NOTE: There were 58151 observations read from the data set D10.DTBD104.
NOTE: There were 130295 observations read from the data set D10.EXPD101.
NOTE: There were 130265 observations read from the data set D10.EXPD102.
NOTE: There were 119867 observations read from the data set D10.EXPD103.
NOTE: There were 119285 observations read from the data set D10.EXPD104.
NOTE: The data set WORK.EXPEND has 741376 observations and 3 variables.
NOTE: DATA statement used (Total process time):
      real time           1.65 seconds
      cpu time            0.67 seconds


577
578
579
580  DATA PUBFILE (KEEP = NEWID INCLASS UCC RCOST1-RCOST45);
581    MERGE FMLY   (IN = INFAM)
582          EXPEND (IN = INEXP);
583    BY NEWID;
584    IF INEXP AND INFAM;
585
586    IF COST = .  THEN
587       COST = 0;
588
589       ARRAY REPS_A(45) REPWT1-REPWT45;
590       ARRAY REPS_B(45) RCOST1-RCOST45;
591
592       DO i = 1 TO 45;
593         IF REPS_A(i)> 0
594            THEN REPS_B(i) = (REPS_A(i) * COST);
595            ELSE REPS_B(i) = 0;
596       END;
597       /* MERGE FMLY FILE WEIGHTS AND CHARACTERISTICS WITH EXPN/DTAB COSTS */
598       /* MULTIPLY COSTS BY WEIGHTS TO DERIVE WEIGHTED COSTS               */
599  RUN;

NOTE: There were 14296 observations read from the data set WORK.FMLY.
NOTE: There were 741376 observations read from the data set WORK.EXPEND.
NOTE: The data set WORK.PUBFILE has 741376 observations and 48 variables.
NOTE: DATA statement used (Total process time):
      real time           12.31 seconds
      cpu time            4.03 seconds


600
601
602    /*************************************************************************/
603    /* STEP3: CALCULATE POPULATIONS                                          */
604    /* --------------------------------------------------------------------- */
605    /* 1 SUM ALL 45 WEIGHT VARIABLES TO DERIVE REPLICATE POPULATIONS         */
606    /* 2 FORMAT FOR CORRECT COLUMN CLASSIFICATIONS                           */
607    /*************************************************************************/
608
609
610  PROC SUMMARY NWAY DATA=FMLY;
```

Reads in all DTAB income data and EXPN expenditure data.

Newid is the consumer unit code. UCC is a code that represents the type of expenditure variable. Cost is the value that corresponds to the UCC code.

Merges the FMLY and EXPEND data sets together and changes missing cost values to zero.

Weights the cost values by the 44 replicate weights and full sample weight. RCOST1-RCOST45 represents the weighted costs for each expenditure.

```
611    CLASS INCLASS / MLF;
612    VAR REPWT1-REPWT45;
613    FORMAT INCLASS $INC.;
614    OUTPUT OUT = POP (DROP = _TYPE_ _FREQ_) SUM = RPOP1-RPOP45;
615    /* SUMS WEIGHTS TO CREATE POPULATIONS PER REPLICATE */
616    /* FORMATS TO CORRECT COLUMN CLASSIFICATIONS        */
617  RUN;

NOTE: There were 14296 observations read from the data set WORK.FMLY.
NOTE: The data set WORK.POP has 10 observations and 46 variables.
NOTE: PROCEDURE SUMMARY used (Total process time):
      real time           0.20 seconds
      cpu time            0.06 seconds


618
619
620
621    /*************************************************************************/
622    /* STEP4: CALCULATE WEIGHTED AGGREGATE EXPENDITURES                      */
623    /* --------------------------------------------------------------------- */
624    /* 1 SUM THE 45 REPLICATE WEIGHTED EXPENDITURES TO DERIVE AGGREGATES     */
625    /* 2 FORMAT FOR CORRECT COLUMN CLASSIFICATIONS AND AGGREGATION SCHEME    */
626    /*************************************************************************/
627
628
629  PROC SUMMARY NWAY DATA=PUBFILE SUMSIZE=MAX COMPLETETYPES;
630    CLASS UCC INCLASS / MLF;
631    VAR RCOST1-RCOST45;
632    FORMAT UCC $AGGFMT. INCLASS $INC.;
633     OUTPUT OUT=AGG (DROP= _TYPE_ _FREQ_  RENAME=(UCC=LINE))
634     SUM = RCOST1-RCOST45;
635    /* SUMS WEIGHTED COSTS PER REPLICATE TO GET AGGREGATES */
636    /* FORMATS INCOME TO CREATE COMPLETE REPORTING COLUMN  */
637    /* FORMATS EXPENDITURES TO CORRECT AGGREGATION SCHEME  */
638  RUN;

NOTE: There were 741376 observations read from the data set WORK.PUBFILE.
NOTE: The data set WORK.AGG has 4680 observations and 47 variables.
NOTE: PROCEDURE SUMMARY used (Total process time):
      real time           8.34 seconds
      cpu time            9.54 seconds


639
640
641
642    /*************************************************************************/
643    /* STEP5: CALCULTATE MEAN EXPENDITURES                                   */
644    /* --------------------------------------------------------------------- */
645    /* 1 READ IN POPULATIONS AND LOAD INTO MEMORY USING A 2 DIMENSIONAL ARRAY  */
646    /*   POPULATIONS ARE ASSOCIATED BY INCLASS(i), AND REPLICATE(j)          */
647    /* 2 READ IN AGGREGATE EXPENDITURES FROM AGG DATASET                     */
648    /*   CALCULATE MEANS BY DIVIDING AGGREGATES BY CORRECT SOURCE POPULATIONS  */
649    /* 4 CALCULATE STANDARD ERRORS USING REPLICATE FORMULA                   */
650    /*************************************************************************/
651
652
653  DATA TAB1 (KEEP = LINE MEAN SE);
654
655    /* READS IN POP DATASET. _TEMPORARY_ LOADS POPULATIONS INTO SYSTEM MEMORY  */
656    ARRAY POP{01:10,45} _TEMPORARY_;
657    IF _N_ = 1 THEN DO i = 1 TO 10;
658      SET POP;
659      ARRAY REPS(45) RPOP1-RPOP45;
660        DO j = 1 TO 45;
661          POP{INCLASS,j} = REPS(j);
662        END;
663      END;
664
665    /* READS IN AGG DATASET AND CALCULATES MEANS BY DIVIDING BY POPULATIONS  */
```

The weights in the FMLY file are summed to create replicate populations and the full US population for each income class. Replicate populations (Repwt1-Repwt44) and the US population (Repwt45) are used as the denominator in means estimation.

Weighted costs are summed and formatted into income classes and by the aggregation scheme of the stub file. These aggregate expenditures will become the numerator in means estimation.

This data step calculates means and standard errors:

Lines 665-662 read in the column populations and stores them into temporary memory. Populations in memory are associated with INCLASS(i), and REPLICATE(j).

```
666    SET AGG (KEEP = LINE INCLASS RCOST1-RCOST45);
667      ARRAY AGGS(45) RCOST1-RCOST45;
668      ARRAY AVGS(45) MEAN1-MEAN44 MEAN;
669        DO k = 1 TO 45;
670          IF AGGS(k) = . THEN AGGS(k) = 0;
671          AVGS(k) = AGGS(k) / POP{INCLASS,k};
672        END;
673
674    /* CALCULATES STANDARD ERRORS USING REPLICATE FORMULA  */
675    ARRAY RMNS(44) MEAN1-MEAN44;
676    ARRAY DIFF(44) DIFF1-DIFF44;
677      DO n = 1 TO 44;
678        DIFF(n) = (RMNS(n) - MEAN)**2;
679      END;
680    SE = SQRT((1/44)*SUM(OF DIFF(*)));
681  RUN;

NOTE: Character values have been converted to numeric values at the places given by:
(Line):(Column).
      661:13   671:33
NOTE: There were 10 observations read from the data set WORK.POP.
NOTE: There were 4680 observations read from the data set WORK.AGG.
NOTE: The data set WORK.TAB1 has 4680 observations and 3 variables.
NOTE: DATA statement used (Total process time):
      real time           0.14 seconds
      cpu time            0.07 seconds


682
683
684
685    /*************************************************************************/
686    /* STEP6: TABULATE EXPENDITURES                                          */
687    /* --------------------------------------------------------------------- */
688    /* 1 ARRANGE DATA INTO TABULAR FORM                                      */
689    /* 2 SET OUT DIARY POPULATIONS FOR POPULATION LINE ITEM                  */
690    /* 3 INSERT POPULATION LINE INTO TABLE                                   */
691    /* 4 INSERT ZERO EXPENDITURE LINE ITEMS INTO TABLE FOR COMPLETENESS      */
692    /*************************************************************************/
693
694
695  PROC TRANSPOSE DATA=TAB1 OUT=TAB2
696    NAME = ESTIMATE PREFIX = INCLASS;
697    BY LINE;
698    VAR MEAN SE;
699    /*ARRANGE DATA INTO TABULAR FORM */
700  RUN;

NOTE: There were 4680 observations read from the data set WORK.TAB1.
NOTE: The data set WORK.TAB2 has 936 observations and 12 variables.
NOTE: PROCEDURE TRANSPOSE used (Total process time):
      real time           0.03 seconds
      cpu time            0.00 seconds


701
702
703  PROC TRANSPOSE DATA=POP (KEEP = RPOP45) OUT=CUS
704    NAME = LINE PREFIX = INCLASS;
705    VAR RPOP45;
706    /* SET ASIDE POPULATIONS FROM DIARY */
707  RUN;

NOTE: There were 10 observations read from the data set WORK.POP.
NOTE: The data set WORK.CUS has 1 observations and 11 variables.
NOTE: PROCEDURE TRANSPOSE used (Total process time):
      real time           0.01 seconds
      cpu time            0.00 seconds


708
```

Line 667 reads in the aggregated expenditures.

Lines 668-672 calculate means by dividing the aggregate expenditures by the appropriate populations in memory as determined by INCLASS and REPLICATE.

Lines 675-680 calculate standard errors using the replicate weight formula.

Arranges output for tabulation. This will give a rough expenditure table.

All populations are put into dataset POP. A special dataset, CUS, is created specifically for inserting the full US population into the output.

```
709
710  DATA TAB3;
711    SET CUS TAB2;
712    IF LINE = 'RPOP45' THEN DO;
713      LINE = '100001';
714      ESTIMATE = 'N';
715      END;
716    /* INSERT POPULATION LINE ITEM INTO TABLE AND ASSIGN LINE NUMBER */
717  RUN;

NOTE: There were 1 observations read from the data set WORK.CUS.
NOTE: There were 936 observations read from the data set WORK.TAB2.
NOTE: The data set WORK.TAB3 has 937 observations and 12 variables.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


718
719
720  DATA TAB;
721    MERGE TAB3 STUBFILE;
722    BY LINE;
723      IF LINE NE '100001' THEN DO;
724        IF SURVEY = 'S' THEN DELETE;
725      END;
726      ARRAY CNTRL(10) INCLASS1-INCLASS10;
727        DO i = 1 TO 10;
728          IF CNTRL(i) = . THEN CNTRL(i) = 0;
729          IF SUM(OF CNTRL(*)) = 0 THEN ESTIMATE = 'MEAN';
730        END;
731
732      IF GROUP IN ('CUCHARS' 'INCOME') THEN DO;
733        IF LAG(LINE) = LINE THEN DELETE;
734      END;
735    /* MERGE STUBFILE BACK INTO TABLE TO INSERT EXPENDITURE LINES */
736    /* THAT HAD ZERO EXPENDITURES FOR THE YEAR                    */
737  RUN;

NOTE: There were 937 observations read from the data set WORK.TAB3.
NOTE: There were 480 observations read from the data set WORK.STUBFILE.
NOTE: The data set WORK.TAB has 872 observations and 20 variables.
NOTE: DATA statement used (Total process time):
      real time           0.03 seconds
      cpu time            0.03 seconds


738
739
740  PROC TABULATE DATA=TAB;
741    CLASS LINE / GROUPINTERNAL ORDER=DATA;
742    CLASS ESTIMATE;
743    VAR INCLASS1-INCLASS10;
744    FORMAT LINE $LBLFMT.;
745
746      TABLE (LINE * ESTIMATE), (INCLASS10 INCLASS1 INCLASS2 INCLASS3 INCLASS4
747                                INCLASS5  INCLASS6 INCLASS7 INCLASS8 INCLASS9)
748      *SUM='' / RTS=25;
749      LABEL ESTIMATE=ESTIMATE LINE=LINE
750            INCLASS1='LESS THAN $5,000'   INCLASS2='$5,000 TO $9,999'
751            INCLASS3='$10,000 TO $14,999' INCLASS4='$15,000 TO $19,999'
752            INCLASS5='$20,000 TO $29,999' INCLASS6='$30,000 TO $39,999'
753            INCLASS7='$40,000 TO $49,999' INCLASS8='$50,000 TO $69,999'
754            INCLASS9='$70,000 AND OVER'   INCLASS10='ALL CONSUMER UNITS';
755      OPTIONS NODATE NOCENTER NONUMBER LS=167 PS=MAX;
756      WHERE LINE NE 'OTHER';
757      TITLE "DIARY EXPENDITURES FOR &YEAR BY INCOME BEFORE TAXES";
758  RUN;

NOTE: There were 870 observations read from the data set WORK.TAB.
      WHERE LINE not = 'OTHER';
```

Population totals per income class are inserted into the output.

This data step further processes data by deleting unwanted table line items and inserting zero expenditure lines for items that are not reported. This is to get the output as close to publication tables as possible.

Tabulate the data. Line numbers are formatted to give titles.

```
NOTE: PROCEDURE TABULATE used (Total process time):
      real time           0.10 seconds
      cpu time            0.03 seconds
```

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 the sample CUs for two consecutive 1-week periods; the sample is surveyed across a 12-month period.

2) An Interview panel survey in which each CU in the sample is interviewed once every 3 months over five consecutive quarters to obtain a year's worth of data. 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 The U.S. Census Bureau 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. Respondents are not limited to recording expense for these items only.

A Household Characteristics Questionnaire is completed to record demographic and family characteristics data pertaining to age, sex, race, marital status, and CU relationships each CU member. Income information, such as wage, salary, unemployment compensation, child support, and alimony, as well as information on the employment of each CU member age 14 and over is collected. 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 format 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 demographic and characteristics information pertaining to CU 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 CU income, employment and earnings of CU members. These data, along with the other household characteristics information, permit data users to classify sample units for research purposes, and allow BLS to adjust population weights for CUs who do not cooperate in the survey.

IX. DATA COLLECTION AND PROCESSING

In addition to its data collection duties, the U.S. 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 U.S. 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 data processing carried out by the U.S. Census Bureau includes programming the Computer Assisted Personal Interview (CAPI) instrument used to collect household characteristics, keying the expenditure data from the diary questionnaire, clerical data editing, and correcting for inconsistencies in the collected data.

The data collected on household characteristics using CAPI are sent directly to the Census Demographic Surveys Division (DSD). Upon completion of the written questionnaire by respondents, the diaries are sent from the regional offices to the Census National Processing Center (NPC) in Jeffersonville, IN. At the NPC, the expenditure data are keyed and codes are applied. The keyed expenditure data are sent to DSD, where they are merged with the household characteristic data. Inconsistencies and errors in the combined data are identified and corrected.

After clerical processing at the NPC, the data are transmitted to the Census Processing Center in Suitland, MD, where they pass through basic quality checks of control counts, missing values, etc. The data are then electronically transmitted to BLS in Washington, DC.

B. BUREAU OF LABOR STATISTICS ACTIVITIES

Upon receipt from the U.S. Census Bureau, the data undergo a series of computer edits that identify and correct irregularities and inconsistencies. Other adjustments apply appropriate sales taxes and derive CU weights based on BLS specifications. In addition, demographic and work experience items are imputed when missing or invalid. All data changes and imputations are identified with flags on the Interview data base.

Next, 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 reviewing images of the questionnaires. Errors discovered through this procedure are corrected prior to release of the data.

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 characteristic 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 CE 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 (PSUs), which consist of counties (or parts thereof) or groups of counties. The set of sample PSUs used for the 2010 sample is composed of 91 areas. The design classifies the PSUs into four categories:

- 21 "A" certainty PSUs are Metropolitan Statistical Areas (MSA's) with a population greater than 1.5 million.
- 38 "X" PSUs, are medium-sized MSAs.
- 16 "Y" PSUs are nonmetropolitan areas that are included in the CPI.
- 16 "Z" PSUs 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 2010 survey is generated from the 2000 Population Census file. The sampling frame is augmented by new construction permits and by techniques used to eliminate recognized deficiencies in census coverage. All Enumeration Districts (EDs) from the Census that fail to meet the criterion for good addresses for new construction, and all EDs 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.

B. COOPERATION LEVELS

The annual target sample size at the United States level for the Diary Survey is 7,050 participating sample units. To achieve this target the total estimated work load is 12,100 sample units. This allows for refusals, vacancies, or nonexistent sample unit addresses.

Each participating sample unit selected is asked to keep two 1-week diaries. Each diary is treated independently, so response rates are based on twice the number of housing units sampled.

The response rate for the 2010 Diary Survey is 71.5% 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	*Eligible housing unit interviews* Number of potential diaries	Type A nonresponses	Total respondent interviews
25,158	5,170	19,988	5,692	14,296

Type B or C cases are housing units that are vacant, nonexistent, or ineligible for diary placement. 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 CU, the weighting is performed at the CU level. Several factors are involved in determining the weight for each CU 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 both the CU's member composition and 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 CU who own their living quarters. 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 as the CU may represent a different number of CUs with similar characteristics.

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. A CU'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, CUs 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 CUs 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 CU 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. CUs 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 non-institutional 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's) 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 MSA. The general concept of an MSA is one of a large population nucleus together with adjacent communities that 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 and Supplemental Security income, unemployment compensation, workmen'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 taxes, and other taxes.

Geographic regions

CUs are classified by region according to the address at which they reside during the time of 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

*L denotes UCCs that could have negative values.

An underlined UCC represents either a new UCC or a deleted UCC. Please note that new UCCs may not be represented in all quarters. The quarter in which the addition (deletion) occurs is denoted by a leading superscript directly prior to the UCC code. For example, $^{N(D)071}$(UCC) identifies a new (deleted) UCC beginning in Q071.

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
170531	Other noncarbonated beverage/ice
170532	Bottled water
170533	Sports Drinks
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
180720	Vitamin supplements
190111	Lunch at Fast Food
190112	Lunch at Full Service
190113	Lunch at Vending Machine
190114	Lunch at Employer
190115	Lunch at Board
190116	Lunch at Catered Affairs
190211	Dinner at Fast Food
190212	Dinner at Full Service
190213	Dinner at Vending Machine
190214	Dinner at Employer
190215	Dinner at Board
190216	Dinner at Catered Affairs
190311	Snacks at Fast Food
190312	Snacks at Full Service
190313	Snacks at Vend Machine
190314	Snacks at Employer
190315	Snacks at Board
190316	Snacks at Catered Affairs
190321	Breakfast at Fast Food
190322	Breakfast at Full Service
190323	Breakfast at Vending Machine
190324	Breakfast at Employer
190325	Breakfast at Board
190326	Breakfast at Catered Affairs
190911	Board at Fast Food
190912	Board at Full Service
190913	Board at Vending Machine
190914	Board at Employer
190915	Board
190916	Board at Catered Affairs
190921	Catered Affairs at Fast Food
190922	Catered Affairs at Full Service

190923	Catered Affairs at Vending Machine
190924	Catered Affairs at Employer
190925	Catered Affairs at Board
190926	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
200511	Beer at Fast Food
200512	Beer at Full Service
200513	Beer at Vending Machine
200514	Beer at Employer
200515	Beer at Board
200516	Beer at Catered Affairs
200521	Wine at Fast Food
200522	Wine at Full Service
200523	Wine at Vending Machine
200524	Wine at Employer
200525	Wine at Board
200526	Wine at Catered Affairs
200531	Alcoholic Beverage Excluding Beer/Wine Fast Food
200532	Alcoholic Beverage Excluding Beer/Wine Full Service
200533	Alcoholic Beverage Excluding Beer/Wine Vending Machine
200534	Alcoholic Beverage Excluding Beer/Wine at Employer
200535	Alcoholic Beverage Excluding Beer/Wine at Board
200536	Alcoholic Beverage Excluding Beer/Wine Catered Affairs
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
220400	Purchase of property or real estate
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	Cable/Satellite/Com Antenna Serv
270410	Garbage, trash collection
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
310140	Televisions
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
310315	Digital media players and recorders
310320	Sound components, component systems, amplifiers, receivers, turn tables, tape decks, tuners, stereos, speakers, and compact disc sound systems
310241	Streaming Video Files
310242	Downloading Video Files
310314	Digital Audio Players

310331	Miscellaneous sound equipment
310332	Sound equipment accessories
310334	Satellite dishes
310335	Miscellaneous video equipment
310340	Records, CDs, and Audio Tapes
310351	Streaming Audio Files
310352	Downloading Audio Files
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
320220	Lamps and other lighting fixtures
320232	Telephones and accessories
320233	Clocks and other household decorative items
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, refurnishing, 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
360513	Men’s pants and shorts
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
370314	Boys’ pants and shorts
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
380333	Women’s pants and shorts
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
390223	Girls' pants and shorts
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 freshener, etc...
480213	Battery replacement, floormats, seatcovers, 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
500110	Vehicle insurance
520110	State or local vehicle registration
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 or electronic toll passes
520550	Towing charges
520560	Global Positioning Services
520901	Docking and landing fees for boats and planes, boat ramp fees
520902	Rental of motorcycle, motor scooters, moped, etc., 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, gauze, 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
550410	Nonprescription vitamins
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
600903	Global Positioning System Devices
610110	Toys, games, hobbies, tricycles, and battery powered riders
610120	Playground equipment
610130	Musical instruments and accessories
610140	Stamp And Coin Collecting
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, passport fees
620912	Rental of video cassettes, tapes, and discs
620913	Coin-operated pinball/electronic video games
620915	Sport vehicle rental
620925	Lotteries and Parimutuel Losses
620926	Miscellaneous Fees
620930	Online Entertainment Services
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 eye make-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
680904	Dating Services
690110	Computers for non-business use, hardware and software excluding video games
690114	Computer information services
690115	Personal Digital Assistants
690116	Internet Services Away From Home
690210	Telephone answering devices
690230	Typewriters and other office machines for non-business use
999000	Home ownership expense not specified
999900	Taxes not specified

NOTE: The following lists the UCCs necessary to derive expenditures for these “food away” items:

[1] for LUNCH
190111, 190112, 190113, 190114, 190115, 190116

[2] for DINNER
190211, 190212, 190213, 190214, 190215, 190216

[3] for SNACKS

190311, 190312, 190313, 190314, 190315, 190316

[4] for BREAKFAST
190321, 190322, 190323, 190324, 190325, 190326

[5] for CATERED AFFAIRS
190921, 190922,190923, 90924, 190925, 190926

[6] for BOARD
190911, 190912, 190913,190914, 190915, 190916

[7] for BEER
200511, 200512, 200513, 200514, 200515, 200516

[8] for WINE
200521, 200522, 200523, 200524, 200525, 200526

[9] for ALCOHLIC BEVERAGES, EXCL. BEER AND WINE
200531, 200532, 200533, 200534, 200535, 200536

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
	950002	Federal income tax (deducted)
	950003	Additional federal income tax (paid)
*L	950001	Federal income tax refunds
	950012	State/local income tax (deducted)
	950013	Additional state/local income tax (paid)
*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 UCCs contain values of 100 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. These UCCs 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 Dstub file in the Programs folder on the CD shows the UCC aggregation used in the sample program.

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. The diary data dictionary contains detailed descriptions of these variables arranged on a functional basis.

A. FMLY FILE

Diary: FMLY

Variable Name	Start Position	Variable Name	Start Position	Variable Name	Start Position
NEWID	1	FINCAFTX	130	LUMPX_	468
ADDFEDX	9	FINC_FTX	138	MARITAL1	469
ADDFEDX_	17	FINCBEFX	139	MARI_AL1	470
ADDOTHX	18	FINC_EFX	147	NO_EARNR	471
ADDOTHX_	26	FINLWT21	148	NO_E_RNR	473
ADDSTAX	27	FIRAX	159	OCCEXPNX	483
ADDSTAX_	35	FIRAX_	167	OCCE_PNX	491
AGE_REF	36	FJSSDEDX	168	OCCULIS2	492
AGE_REF_	38	FJSS_EDX	176	OCCU_IS2	494
AGE2	39	FPVTX	177	OTHINX	499
AGE2_	41	FPVTX_	185	OTHINX_	507
BLS_URBN	42	FREEMLX	186	OTHRECX	508
CUTENURE	43	FREEMLX_	194	OTHRECX_	516
CUTE_URE	44	FRRX	195	OTHREFX	517
DESCRIP	45	FRRX_	203	OTHREFX_	525
DESCRIP_	47	FS_MTHI	348	OTHRNTX	526
DIVX	48	FS_MTHI_	350	OTHRNTX_	534
DIVX_	56	FSS_RRX	351	PENSIONX	535
EARNCOMP	57	FSS_RRX_	359	PENS_ONX	543
EARN_OMP	58	FSTATXX	360	PERSLT18	544
EDUC_REF	68	FSTATXX_	368	PERS_T18	546
EDUC0REF	70	FSUPPX	369	PERSOT64	547
EDUCA2	71	FSUPPX_	377	PERS_T64	549
EDUCA2_	73	FWAGEX	378	PERSTAX	550
EMPLTYP1	74	FWAGEX_	386	PERSTAX_	558
EMPL_YP1	75	HRSPRWK1	387	OCCULIS1	561
EMPLTYP2	76	HRSP_WK1	390	OCCU_IS1	563
EMPL_YP2	77	HRSPRWK2	391	POPSIZE	564
FAM_SIZE	78	HRSP_WK2	394	PTAXREFX	565
FAM__IZE	80	INSREFX	405	PTAX_EFX	573
FAM_TYPE	81	INSREFX_	413	RACE2	574
FAM__YPE	82	INTX	414	RACE2_	575
FBSNSX	83	INTX_	422	REC_FS	576
FBSNSX_	91	JFS_AMT	423	REC_FS_	577
FD_STMPS	92	JFS_AMT_	431	REF_RACE	578
FD_S_MPS	93	JGRCFDMV	432	REF__ACE	579
FEDREFX	94	JGRC_DMV	438	REGION	580
FEDREFX_	102	JGRCFDWK	439	RESPSTAT	582
FFARMX	103	JGRC_DWK	445	RESP_TAT	583

Diary: FMLY

Variable Name	Start Position	Variable Name	Start Position	Variable Name	Start Position
SEX_REF_	603	WTREP11	797	MILKPROD	1291
SEX2	604	WTREP12	808	OTHDAIRY	1303
SEX2_	605	WTREP13	819	FRSHFRUT	1315
SMSASTAT	606	WTREP14	830	FRSHVEG	1327
SSREFX	607	WTREP15	841	PROCFRUT	1339
SSREFX_	615	WTREP16	852	PROCVEG	1351
STATREFX	616	WTREP17	863	SWEETS	1363
STAT_EFX	624	WTREP18	874	NONALBEV	1375
STRTDAY	625	WTREP19	885	OILS	1387
STRTMNTH	627	WTREP20	896	MISCFOOD	1399
STRTYEAR	629	WTREP21	907	FOODAWAY	1411
TAXPROPX	633	WTREP22	918	ALCBEV	1423
TAXP_OPX	641	WTREP23	929	SMOKSUPP	1435
TYPOWND	642	WTREP24	940	PET_FOOD	1447
TYPOWND_	643	WTREP25	951	PERSPROD	1459
UNEMPX	644	WTREP26	962	PERSSERV	1471
UNEMPX_	652	WTREP27	973	DRUGSUPP	1483
VEHQ	653	WTREP28	984	HOUSKEEP	1495
VEHQ_	655	WTREP29	995	HH_CU_Q	1507
WEEKI	656	WTREP30	1006	HH_CU_Q_	1509
WEEKI_	657	WTREP31	1017	HHID	1510
WEEKN	658	WTREP32	1028	HHID_	1513
WELFRX	659	WTREP33	1039	CHILDAGE	1514
WELFRX_	667	WTREP34	1050	CHIL_AGE	1515
WHYNWRK1	668	WTREP35	1061	INCLASS	1516
WHYN_RK1	669	WTREP36	1072	STATE	1518
WHYNWRK2	670	WTREP37	1083	CHDOTHX	1521
WHYN_RK2	671	WTREP38	1094	CHDOTHX_	1529
WK_WRKD1	672	WTREP39	1105	ALIOTHX	1530
WK_W_KD1	674	WTREP40	1116	ALIOTHX_	1538
WK_WRKD2	675	WTREP41	1127	CHDLMPX	1539
WK_W_KD2	677	WTREP42	1138	CHDLMPX_	1547
WRKRSX	678	WTREP43	1149	INC_RANK	1559
WRKRSX_	686	WTREP44	1160	INC__ANK	1568
WTREP01	687	FOODTOT	1171	CUID	1569
WTREP02	698	FOODHOME	1183	HORREF1	1576
WTREP03	709	CEREAL	1195	HORREF1_	1577
WTREP04	720	BAKEPROD	1207	HORREF2	1578
WTREP05	731	BEEF	1219	HORREF2_	1579
WTREP06	742	PORK	1231	ALIOTHXM	1580
WTREP07	753	OTHMEAT	1243	ALIO_HXM	1590
WTREP08	764	POULTRY	1255	ALIOTHX1	1591
WTREP09	775	SEAFOOD	1267	ALIOTHX2	1599
WTREP10	786	EGGS	1279	ALIOTHX3	1607

Diary: FMLY

Variable Name	Start Position	Variable Name	Start Position	Variable Name	Start Position
ALIOTHX4	1615	FINCAFTM	1922	FSTATXX1	2215
ALIOTHX5	1623	FINC_FTM	1933	FSTATXX2	2223
ALIOTHXI	1631	FINCAFT1	1934	FSTATXX3	2231
CHDOTHXM	1634	FINCAFT2	1943	FSTATXX4	2239
CHDO_HXM	1644	FINCAFT3	1952	FSTATXX5	2247
CHDOTHX1	1645	FINCAFT4	1961	FSUPPXM	2255
CHDOTHX2	1653	FINCAFT5	1970	FSUPPXM_	2265
CHDOTHX3	1661	FINCBEFM	1979	FSUPPX1	2266
CHDOTHX4	1669	FINC_EFM	1990	FSUPPX2	2274
CHDOTHX5	1677	FINCBEF1	1991	FSUPPX3	2282
CHDOTHXI	1685	FINCBEF2	2000	FSUPPX4	2290
DIVXM	1688	FINCBEF3	2009	FSUPPX5	2298
DIVXM_	1698	FINCBEF4	2018	FSUPPXI	2306
DIVX1	1699	FINCBEF5	2027	FWAGEXM	2309
DIVX2	1707	FINCBEFI	2036	FWAGEXM_	2319
DIVX3	1715	FJSSDEDM	2039	FWAGEX1	2320
DIVX4	1723	FJSS_EDM	2049	FWAGEX2	2328
DIVX5	1731	FJSSDED1	2050	FWAGEX3	2336
DIVXI	1739	FJSSDED2	2058	FWAGEX4	2344
FBSNSXM	1742	FJSSDED3	2066	FWAGEX5	2352
FBSNSXM_	1753	FJSSDED4	2074	FWAGEXI	2360
FBSNSX1	1754	FJSSDED5	2082	INC_RNKM	2363
FBSNSX2	1763	FPVTXM	2090	INC__NKM	2372
FBSNSX3	1772	FPVTXM_	2098	INC_RNK1	2373
FBSNSX4	1781	FRRXM	2099	INC_RNK2	2382
FBSNSX5	1790	FRRXM_	2107	INC_RNK3	2391
FBSNSXI	1799	FS_AMTXM	2108	INC_RNK4	2400
FFARMXM	1802	FS_A_TXM	2116	INC_RNK5	2409
FFARMXM_	1813	FS_AMTX1	2117	INTXM	2418
FFARMX1	1814	FS_AMTX2	2123	INTXM_	2428
FFARMX2	1823	FS_AMTX3	2129	INTX1	2429
FFARMX3	1832	FS_AMTX4	2135	INTX2	2437
FFARMX4	1841	FS_AMTX5	2141	INTX3	2445
FFARMX5	1850	FS_AMTXI	2147	INTX4	2453
FFARMXI	1859	FSS_RRXM	2150	INTX5	2461
FFEDTXXM	1862	FSS__RXM	2160	INTXI	2469
FFED_XXM	1872	FSS_RRX1	2161	JFS_AMTM	2472
FFEDTXX1	1873	FSS_RRX2	2169	JFS__MTM	2480
FFEDTXX2	1881	FSS_RRX3	2177	JFS_AMT1	2481
FFEDTXX3	1889	FSS_RRX4	2185	JFS_AMT2	2487
FFEDTXX4	1897	FSS_RRX5	2193	JFS_AMT3	2493
FFEDTXX5	1905	FSS_RRXI	2201	JFS_AMT4	2499
FGVXM	1913	FSTATXXM	2204	JFS_AMT5	2505
FGVXM_	1921	FSTA_XXM	2214	OTHINXM	2511

Diary: FMLY

Variable Name	Start Position	Variable Name	Start Position	Variable Name	Start Position
OTHINXM_	2521	UNEMPX5	2824	OTHINB_	3006
OTHINX1	2522	UNEMPXI	2830	OTHINBX	3007
OTHINX2	2530	WELFRXM	2833	OTHINBX_	3013
OTHINX3	2538	WELFRXM_	2843	OTHLOSSB	3014
OTHINX4	2546	WELFRX1	2844	OTHL_SSB	3016
OTHINX5	2554	WELFRX2	2852	OTHLOSBX	3017
OTHINXI	2562	WELFRX3	2860	OTHL_SBX	3023
OTHRNTXM	2565	WELFRX4	2868	PNSIONB	3024
OTHR_TXM	2576	WELFRX5	2876	PNSIONB_	3026
OTHRNTX1	2577	WELFRXI	2884	PNSIONBX	3027
OTHRNTX2	2586	WRKRSXM	2887	PNSI_NBX	3033
OTHRNTX3	2595	WRKRSXM_	2897	ROOMLOSB	3034
OTHRNTX4	2604	WRKRSX1	2898	ROOM_OSB	3036
OTHRNTX5	2613	WRKRSX2	2906	ROOMLSBX	3037
OTHRNTXI	2622	WRKRSX3	2914	ROOM_SBX	3043
PENSIONM	2625	WRKRSX4	2922	SALEB	3044
PENS_ONM	2635	WRKRSX5	2930	SALEB_	3046
PENSION1	2636	WRKRSXI	2938	SALEBX	3047
PENSION2	2644	PICKCODE	2941	SALEBX_	3053
PENSION3	2652	ALIOTHB	2944	UNEMPB	3054
PENSION4	2660	ALIOTHB_	2946	UNEMPB_	3056
PENSION5	2668	ALIOTHBX	2947	UNEMPBX	3057
PENSIONI	2676	ALIO_HBX	2953	UNEMPBX_	3063
PERSTAXM	2679	CHDLMPB	2954	WELFRB	3064
PERS_AXM	2690	CHDLMPB_	2956	WELFRB_	3066
PERSTAX1	2691	CHDLMPBX	2957	WELFRBX	3067
PERSTAX2	2700	CHDL_PBX	2963	WELFRBX_	3073
PERSTAX3	2709	CHDOTHB	2964	WRKRSB	3074
PERSTAX4	2718	CHDOTHB_	2966	WRKRSB_	3076
PERSTAX5	2727	CHDOTHBX	2967	WRKRSBX	3077
ROOMXM	2743	CHDO_HBX	2973	WRKRSBX_	3083
ROOMXM_	2752	DIVB	2974	PSU	3084
ROOMX1	2753	DIVB_	2976	POVLEVCY	3088
ROOMX2	2760	DIVBX	2977	POVLEVPY	3096
ROOMX3	2767	DIVBX_	2983	POV_CY	3104
ROOMX4	2774	INTB	2984	POV_PY	3105
ROOMX5	2781	INTB_	2986	POV_CYM	3106
ROOMXI	2788	INTBX	2987	POV_CY1	3107
UNEMPXM	2791	INTBX_	2993	POV_CY2	3108
UNEMPXM_	2799	LUMPB	2994	POV_CY3	3109
UNEMPX1	2800	LUMPB_	2996	POV_CY4	3110
UNEMPX2	2806	LUMPBX	2997	POV_CY5	3111
UNEMPX3	2812	LUMPBX_	3003	POV_PYM	3112
UNEMPX4	2818	OTHINB	3004	POV_PY1	3113

Diary: FMLY

	Variable Name	Start Position
	POV_PY2	3114
	POV_PY3	3115
	POV_PY4	3116
	POV_PY5	3117
* N(Y091)	HISP_REF	3118
* N(Y091)	HISP2	3119

Diary: MEMB

Variable Name	Start Position	Variable Name	Start Position	Variable Name	Start Position
NEWID	1	RRX	153	RC_B_ACK	257
AGE	9	RRX_	161	RC_NATAM	258
AGE_	11	SCHLNCHQ	162	RC_N_TAM	259
ANFEDTXX	12	SCHL_CHQ	164	RC_ASIAN	260
ANFE_TXX	20	SCHLNCHX	165	RC_A_IAN	261
ANGVX	21	SCHL_CHX	173	RC_PACIL	262
ANGVX_	29	SEX	174	RC_P_CIL	263
ANPVTX	30	SLFEMPSS	176	RC_OTHER	264
ANPVTX_	38	SLFE_PSS	182	RC_O_HER	265
ANRRX	39	SS_RRX	183	RC_DK	266
ANRRX_	47	SS_RRX_	191	RC_DK_	267
ANSTATXX	48	STA_SUPP	192	ANFEDTXM	268
ANST_TXX	56	STA__UPP	193	ANFE_TXM	276
ANYRAIL	57	STATXX	194	ANGVXM	277
ANYRAIL_	58	STATXX_	202	ANGVXM_	285
ANYSSINC	59	SUPPX	203	ANPVTXM	286
ANYS_INC	60	SUPPX_	211	ANPVTXM_	294
BSNSX	61	US_SUPP	212	ANRRXM	295
BSNSX_	69	US_SUPP_	213	ANRRXM_	303
CU_CODE1	70	WAGEX	214	ANSTATXM	304
EDUCA	72	WAGEX_	222	ANST_TXM	312
EDUCA_	74	WHYNOWRK	223	BSNSXM	313
EMPLTYPE	75	WHYN_WRK	224	BSNSXM_	324
EMPL_YPE	76	WKS_WRKD	225	BSNSX1	325
FARMX	77	WKS__RKD	227	BSNSX2	334
FARMX_	85	SS_RRQ	228	BSNSX3	343
FEDTXX	86	SS_RRQ_	232	BSNSX4	352
FEDTXX_	94	SOCRRX	233	BSNSX5	361
GROSPAYX	95	SOCRRX_	241	BSNSXI	370
GROS_AYX	103	ARM_FORC	242	FARMXM	373
GVX	104	ARM__ORC	243	FARMXM_	384
GVX_	112	IN_COLL	244	FARMX1	385
HRSPERWK	113	IN_COLL_	245	FARMX2	394
HRSP_RWK	116	MEDICARE	246	FARMX3	403
IRAX	117	MEDI_ARE	247	FARMX4	412
IRAX_	125	PAYPERD	248	FARMX5	421
JSSDEDX	126	PAYPERD_	249	FARMXI	430
JSSDEDX_	132	HORIGIN	250	JSSDEDXM	433
MARITAL	133	HISPANIC	251	JSSD_DXM	441
MEMBNO	135	HISP_NIC	252	JSSDEDX1	442
OCCULIST	137	MEMBRACE	253	JSSDEDX2	448
OCCU_IST	139	RC_WHITE	254	JSSDEDX3	454
PVTX	142	RC_W_ITE	255	JSSDEDX4	460
PVTX_	150	RC_BLACK	256	JSSDEDX5	466

Diary: MEMB

Variable Name	Start Position	Variable Name	Start Position
SLFEMPSM	472	FARMBX	737
SLFE_PSM	480	FARMBX_	743
SLFEMPS1	481	SS_RRB	744
SLFEMPS2	487	SS_RRB_	746
SLFEMPS3	493	SS_RRBX	747
SLFEMPS4	499	SS_RRBX_	753
SLFEMPS5	505	SUPPB	754
SOCRRXM	511	SUPPB_	756
SOCRRXM_	521	SUPPBX	757
SOCRRX1	522	SUPPBX_	763
SOCRRX2	530	WAGEB	764
SOCRRX3	538	WAGEB_	766
SOCRRX4	546	WAGEBX	767
SOCRRX5	554	WAGEBX_	773
SS_RRXM	562	ASIAN	774
SS_RRXM_	572	ASIAN_	775
SS_RRX1	573		
SS_RRX2	581		
SS_RRX3	589		
SS_RRX4	597		
SS_RRX5	605		
SS_RRXI	613		
SUPPXM	616		
SUPPXM_	626		
SUPPX1	627		
SUPPX2	635		
SUPPX3	643		
SUPPX4	651		
SUPPX5	659		
SUPPXI	667		
WAGEXM	670		
WAGEXM_	680		
WAGEX1	681		
WAGEX2	689		
WAGEX3	697		
WAGEX4	705		
WAGEX5	713		
WAGEXI	721		
BSNSB	724		
BSNSB_	726		
BSNSBX	727		
BSNSBX_	733		
FARMB	734		
FARMB_	736		

XVI. APPENDIX 5--PUBLICATIONS AND DATA RELEASES FROM THE CONSUMER EXPENDITURE SURVEY

CONSUMER EXPENDITURE SURVEY DATA ON THE INTERNET

CE reports and data tables can be found on-line at http://www.bls.gov/cex/home.htm .
The following One and Two-year Tables of integrated Diary and Interview data are available under the Tables Created by BLS heading:

One Year Tables

Standard Tables from 1984-2010
Expenditure Shares Tables from 1998-2010
Aggregate Expenditure Shares Tables from 1998-2010

Two Year Tables

Cross-Tabulated Tables from 1986-2010
Metropolitan Statistical Area Tables from 1986-2010
Region Tables from 1998-2010
High Income Tables from 1998-2002
Multi-Year Tables for 1984-1992 and 1994-2010

CD-ROMS

CE microdata on CD-ROM are available from the Bureau of Labor Statistics for 1972-73, 1980-81, 1990-91, 1992-93, and for each individual year from 1994-2010. The 1980-81 through 2010 releases contain Interview and Diary data, while the 1972-73 CD includes Interview data only. The 1980-81, 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 CD do not. In addition to the Interview and Diary data, the CDs from 1994-2004 include the complete collection of EXPN files. A 1984-94 "multi-year" CD 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 (1984-present) that are found on the Consumer Expenditure Survey web site (http://www.bls.gov/cex).

More information on the particular CD-ROMs available and the order form can be found on the Consumer Expenditure Survey web site: http://www.bls.gov/cex/csxmicro.htm

XVII. INQUIRIES, SUGGESTIONS, AND COMMENTS

If you have any questions, suggestions, or comments about the survey, the microdata, or its documentation please call (202) 691-6900 or email cexinfo@bls.gov .

Written suggestions and comments should be forwarded to:

Division of Consumer Expenditure Survey
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.