Getting started with Consumer Expenditure Survey (CE) Public-Use Microdata (PUMD)[1]

Division of Consumer Expenditure Survey, Bureau of Labor Statistics[2]

August 30, 2016

[1] Version 3.0

[2] For questions or comments, please contact the CE by phone, (202) 691-6900, or email, CEXINFO@bls.gov.

Contents

1 Introduction 1

1.1 What surveys does CE conduct? 1

1.2 What is CE PUMD? 1

1.3 How is the Getting Started Guide Organized? 2

2 File naming conventions 3

2.1.1 File names and directories 3

2.1.2 What's the "X" about? 4

3 Interview survey data 4

3.1 "Annual" or "5 quarter" files 5

3.2 Paradata 5

3.3 Interview quarterly files 6

3.3.1 FMLI-CU characteristics, income, and summary expenditure variables 6

3.3.2 MEMI-CU member characteristics and income 6

3.3.3 MTBI-Monthly expenditure file classified by UCC 7

3.3.4 ITBI-Consumer Unit monthly income classified by UCC 8

3.3.5 ITII-Consumer Unit monthly imputed income file classified by UCC. 8

3.3.6 NTAXI-Federal and state tax information classified by tax unit 9

3.4 Interview annual files–EXPN 9

3.4.1 Three examples of EXPN files: 10

1 CLA-Clothing and sewing materials 10

Unique records in CLA are defined by the combination of NEWID, SEQNO, and ALCNO 10

2 IHB-Hospitalization and health insurance detailed questions 10

3 VEQ-Vehicle maintenance and repair 11

4 Diary Survey Data 12

4.1 Diary files 12

4.1.1 FMLD-CU characteristics, income, weights, and summary level expenditures 12

4.1.2 MEMD-CU member characteristics and income 13

4.1.3 EXPD-Detailed weekly expenditure file categorized by UCC 13
4.1.4 DTBD-Detailed annual income file categorized by UCC 14
4.1.5 DTID-Consumer Unit imputed income file categorized by UCC. 14
5 Documentation 14

1 Introduction

The *Getting Started Guide* serves to help users of Consumer Expenditure Survey's Public-Use Microdata. The Guide is intended for novice users of Consumer Expenditure Survey (CE) microdata, however, users should be familiar with data structures and basic statistical concepts.

CE is a program of the U.S. Bureau of Labor Statistics which is responsible for conducting surveys to report expenditure, income, and tax statistics about consumer units across the United States. Please visit the Consumer Expenditure Survey website for more information.

1.1 What surveys does CE conduct?

CE collects data through two surveys: the Quarterly Interview Survey (Interview) and the Diary Survey (Diary). The Interview generally tracks consumer units' large expenditures, such as major appliances and cars, while the Diary tracks smaller, everyday expenditures that might be easily forgotten even after a few days such as a cup of coffee. The Interview is conducted quarterly with each consumer unit and the Diary is conducted over two consecutive one-week periods with each respondent.

Please note that while there is some overlap in the expenses recorded in each survey there is no overlap in the samples, so a consumer unit in one survey will not appear in the other. While integrating the data from the samples is possible, the statistical methods involved in doing so are complex. CE offers integrated tables for the expense categories that appear in both.

1.2 What is CE PUMD?

CE provides data in three ways with different levels of detail and complexity:

A. Major pre-defined expenditures, income, and demographic tables by age, race, region, housing, education, occupation, income categories, number of earners, Hispanic or Latino origin, and other aspects. In addition, we provide tables that cross tabulate two characteristics, for example age and income. To use these tables, users do not need statistical knowledge or software. These tables are available in three publications: News releases, published tables, and database.

B. Detailed pre-defined expenditure, income, and demographic tables provide the same data as the tabulations described above with additional detail and the standard error. To use these tables, users need to take into account that some data points have a high standard error that may make the data inadequate for some purposes. To obtain "prepubs" please submit a Request for Consumer Expenditure Survey (CE) Information.

C. Public-use microdata (PUMD) files contain the actual responses to the interview and diary survey. The responses can be used to create and correlate variables. CE adjusts this dataset

in two ways. First, CE estimates missing components, like income, and second, anonymizes the data to protect respondent identities.

To use this dataset, users need to be familiar with statistical concepts, such as standard error, and be proficient with a statistical software package, such as SAS or R. This dataset is available at the PUMD data page.

If your research requires access to our microdata without the alterations made to PUMD you can apply to be an onsite researcher. The applicant must be affiliated with an academic institution or an NGO and be working on strictly statistical research that is in the public interest. Processing time of the application plus preparation of the data requested by the researcher can vary. Applicants also must be U.S. citizens or have a valid visa and social security number and must obtain Title Thirteen (Protection of Confidential Information) clearance through the U.S. Census Bureau. For more information about becoming an onsite researcher, please submit a Request for Consumer Expenditure Survey (CE) Information.

1.3 How is the Getting Started Guide Organized?

This guide is divided into five (5) sections described below: Introduction, File Naming Conventions, The Interview Survey Files, The Diary Survey Files, and Documentation.

Section 1: Introduction
Description of the Getting Started Guide with an overview of CE and the types of available data.

Section 2: File Naming Conventions
An explanation of how files for individual releases are named to aid the user in navigating the database.

Section 3: Interview Survey Files
Description of the file structure of the Interview and the contents of its individual files.

Section 4: Diary Survey Files
Similar to Section 3, this is a description of the file structure of the Diary and the contents of its individual files.

Section 5: Documentation
The Documentation section consists of links to resource material useful for understanding CE methodology and more detailed information about each data release and its associated files.

2 File naming conventions

The naming conventions used for each PUMD release is common to both Interview and Diary data. *YY* stands for the two digit year and *Q* stands for the calendar quarter of the Interview or calendar quarter of the placement of the Diary booklet. Below, the example file directories and names are for SAS datasets. The only difference in the naming convention for other dataset types (i.e., SPSS, STATA, etc.) is the file extension.

2.1.1 File names and directories

Interview quarterly files

FMLI \INTRVW*YY* \FMLI*YYQ*.sas7bdat

MEMI \INTRVW*YY* \MEMI*YYQ*.sas7bdat

MTBI \INTRVW*YY* \MTBI*YYQ*.sas7bdat

ITBI \INTRVW*YY* \ITBI*YYQ*.sas7bdat

ITII \INTRVW*YY* \ITII*YYQ*.sas7bdat

NTAXI \INTRVW*YY* \NTAXI*YYQ*.sas7bdat

Interview annual files

Note: Only selected EXPN files are listed here. All other EXPN files follow the same convention.

CLA \EXPN*YY* \CLA*YY*.sas7bdat

IHB \EXPN*YY* \IHB*YY*.sas7bdat

VEQ \EXPN*YY* \VEQ*YY*.sas7bdat

Diary quarterly files

FMLD \DIARY*YY* \FMLD*YYQ*.sas7bdat

MEMD \DIARY*YY* \MEMD*YYQ*.sas7bdat

EXPD \DIARY*YY* \EXPD*YYQ*.sas7bdat

DTBD \DIARY*YY* \DTBD*YYQ*.sas7bdat

DTID \DIARY*YY* \DTID*YYQ*.sas7bdat

2.1.2 What's the "X" about?

You may have noticed there is an "X" in the name for some of the Interview quarterly files (eg. FMLI*YY*Qx.sas7bdat or MEMI*YY*Qx.sas7bdat). You may have also noticed that this "X" is only in the name of quarter 1 files for the current calendar year. This is because files for the first quarter of any calendar year will appear in two CE PUMD releases:

1. As the "fifth" quarter file in the previous calendar year's release,
2. And as the "first" quarter file in the current calendar year's release.

The "X" signifies that the first quarter file of the current calendar year release is not identical to the fifth quarter file of the previous calendar year release. One reason the files are not identical is because PUMD production processes use the five quarters of data within the release to calculate certain values, such as the critical and topcode values of income and expenditures.

During years in which there was a sample design change, there is a bigger difference in the "X" files. For example, the CE changed its sample design in 2005Q1. When this happens, the CE draws an "overlapping sample" for 2005Q1, i.e., some CUs are for the *old* design, some CUs are for the *new* design, and some are for *both* designs.[3] For the 2004 PUMD release, the CUs in the 2005Q1 files are those for the *old* design. For the 2005 PUMD release, the CUs in the 2005Q1 files are those for the *new* design. CUs that were in both the *old* and *new* designs are contained within the 2005Q1 file for both the 2004 and 2005 PUMD releases.

3 Interview survey data

An Interview "quarter" refers to the calendar quarter in which the interview occurred. For example, any Consumer Unit (CU) interviewed in April, May, or June would have their data stored in the quarter 2 (YYQ2) datasets. During an interview, the CU is asked to report expenditures for a reference period of three months. So, for a CU interviewed in April, their expenditures in the YYQ2 file are for January, February, and March. (This is important to remember when calculating calendar year estimates.) Five quarters of interviews are included in each PUMD release, YYQ1 of (RELEASE YEAR) through YYQ1 of (RELEASE YEAR + 1). There are six types of data files organized by quarter:

1. FMLI - a file with characteristics, income, weights, and summary level expenditures for the CU.
2. MEMI - a file with characteristics and income for each member in the CU.
3. MTBI - a detailed monthly expenditure file categorized by Universal Classification Code (UCC).
4. ITBI - a Consumer Unit monthly income file categorized by UCC.

[3] CUs can be in both the *old* and *new* designs because the designs are based on areas of the United States called Primary Sampling Units (PSU). So if a CU was in a PSU that was in both designs, then they are used in both designs.

5. ITII - a Consumer Unit monthly imputed income file categorized by UCC.[4]

6. NTAXI - a file with federal and state tax information for each tax unit in the CU.[5]

3.1 "Annual" or "5 quarter" files

The Interview EXPN files contain expenditure and non-expenditure information collected directly from all major sections of the quarterly Interview survey. For each PUMD release, there are roughly 50 annual EXPN datasets, with each dataset corresponding to a specific section of the Interview survey (http://www.bls.gov/cex/csxsurveyforms.htm#interview). Because these datasets correspond directly to the Interview survey sections, there are varying reference periods for the expenditures contained within each EXPN dataset, dependent upon the questions within each section. Each annual EXPN dataset is made up of the five quarters of data within each release.

3.2 Paradata

Beginning in 2009, the CE began releasing paradata, which is data about the interview process. Paradata is available in two datasets:

1. FPAR - a file with data about the survey, including timing for each section and whether the respondent used records.

2. MCHI - a file with data about the contact history between the field representative and the respondent, including reasons for interview refusal and time of contact.

Typically, each paradata file has nine quarters of data, starting with YYQ1 from the (RELEASE YEAR - 1) to YYQ1 from the (RELEASE YEAR + 1). Nine quarters are included within each PARADATA file so users can have paradata records across all interviews for a given CU. (A CU completing their last interview during YYQ1 of (RELEASE YEAR) has the potential to have had their first (bounding) interview during YYQ1 of (RELEASE YEAR - 1)).[6]

The Interview Survey is a rotating panel survey in which approximately 12,000 addresses are contacted each calendar quarter of the year for the survey. One-fourth of the addresses that are contacted each quarter are new to the survey. Usable interviews are obtained from approximately 6,900 households at those addresses each quarter of the year. After a housing unit has been in the sample for four consecutive quarters, it is dropped from the survey, and a new address is selected to replace it. Before 2015, the Interview Survey included a preliminary bounding interview, and each CU could be contacted up to five times over five quarters. The bounding interview, which recorded recent major expenditures for comparison with subsequent purchases, was determined to be unnecessary, and was dropped at the beginning of 2015 to save money and reduce respondent

[4] Available beginning in 2004

[5] Available beginning in 2013Q2

[6] With the 2009 PUMD release, historical paradata for 2005-2007 were released. The dataset for 2005 contains 3 quarters of data, while the datasets for 2006 and 2007 contains 4 quarters of data each.

burden and collection costs. See http://www.bls.gov/cex/research_papers/pdf/Recommendation-Regarding-the-Use-of-a-CE-Bounding-Interview.pdf.

3.3 Interview quarterly files

3.3.1 FMLI-CU characteristics, income, and summary expenditure variables

- There is one record per CU.
- Each CU is uniquely identified by NEWID.
- Variables include demographics for the reference person and spouse of reference person, income at the CU level, sample housing unit information, summary level expenditures, and weighting variables.

Note: Not all the variables within the FMLI file are shown in the sample below.

	NEWID	AGE_REF	FAM_SIZE	FINCBTXM	FDAWAYPQ	FDAWAYCQ	FINLWT21
1	2545235	82	2	11807.0	0	0	1980.829
2	2545245	69	3	33773.0	370	0	22347.181
3	2545255	22	1	11000.0	180	0	15272.973
4	2545275	57	1	163793.8	680	0	14774.904
5	2545305	45	2	116227.0	770	0	18557.247
6	2545355	37	1	120000.0	2000	0	16921.900
7	2545365	29	4	31400.0	625	0	15776.683
8	2545375	53	2	200.0	1530	0	15271.262
9	2545385	64	1	60000.0	455	0	16723.973
10	2545405	46	2	84608.2	1540	0	16474.727

3.3.2 MEMI-CU member characteristics and income

- There are multiple records per CU.
- There is one record per member.
- Unique records are defined by the combination of NEWID and MEMBNO.
- Variables include demographics about CU members, member level income, and member relationship status to the reference person.

Note: Not all the variables within the MEMI file are shown in the sample below.

	NEWID	MEMBNO	AGE	CU_CODE	EARNER	EDUCA	SEX	MEMBRACE	SALARYXM
1	2667582	4	21	3	NA	NA	1	1	100
2	2667602	1	26	1	1	2	1	1	745
3	2667602	2	23	0	1	5	2	1	3282
4	2667602	3	21	3	NA	NA	1	1	100
5	2667612	1	22	1	1	5	1	4	871
6	2667612	2	24	6	1	5	1	4	1605
7	2667612	3	57	7	1	4	1	4	817
8	2667612	4	49	7	2	4	2	4	100
9	2667642	1	29	1	1	4	1	1	312
10	2667652	1	28	1	1	7	1	1	100

3.3.3 MTBI-Monthly expenditure file classified by UCC

- There are multiple records per CU.
- There is one record per expenditure per purchase month/year.
- Prior to the 2010 PUMD release, unique records were not identified. From the 2010 release forward, unique records can be identified using a combination of NEWID, EXPNAME, SEQNO, and ALCNO.
- The monthly expenditure amounts are mapped from corresponding EXPN variables.

Note: Not all the variables within the MTBI file are shown below.

	NEWID	UCC	COST	COST_	GIFT	PUBFLAG	REF_MO	REF_YR
1	2554935	210110	232	1	2	2	3	2013
2	2554935	210110	232	1	2	2	2	2013
3	2554935	210110	232	1	2	2	1	2013
4	2545765	210110	600	1	2	2	3	2013
5	2545765	210110	600	1	2	2	2	2013
6	2545765	210110	600	1	2	2	1	2013
7	2562785	210110	500	1	2	2	4	2013
8	2562785	210110	500	1	2	2	3	2013
9	2562785	210110	500	1	2	2	2	2013
10	2560385	210110	660	1	2	2	4	2013

3.3.4 ITBI-Consumer Unit monthly income classified by UCC

- There are multiple records per CU.
- There is one record per income item per month/year.
- Unique records are defined by NEWID, UCC, REFMO, and REFYR.
- The monthly amounts are mapped from the corresponding CU-level imputed income mean values.

Note: This ITBI sample shows ALL variables that are in the full ITBI file.[7]

	NEWID	REFMO	REFYR	UCC	PUBFLAG	VALUE	VALUE_
1	2545235	1	2013	5100	2	520.8333	NA
2	2545235	2	2013	5100	2	520.8333	NA
3	2545235	3	2013	5100	2	520.8333	NA
4	2545235	1	2013	5110	2	520.8333	NA
5	2545235	2	2013	5110	2	520.8333	NA
6	2545235	3	2013	5110	2	520.8333	NA
7	2545235	1	2013	5200	2	2500.0000	NA
8	2545235	2	2013	5200	2	2500.0000	NA
9	2545235	3	2013	5200	2	2500.0000	NA
10	2545235	1	2013	5210	2	2500.0000	NA

3.3.5 ITII-Consumer Unit monthly imputed income file classified by UCC.

- There are multiple records per CU.
- There is one record per income item per month/year.
- Unique records are defined by NEWID, UCC, REFMO, REFYR, and IMPNUM.
- The monthly amounts are mapped from the corresponding CU-level imputed income imputes(5 imputes per income item).
- This file is necessary to calculate variances and standard errors for imputed income.

Note: This ITII sample shows ALL variables that are in the full ITII file.

	NEWID	REFMO	REFYR	UCC	PUBFLAG	VALUE	VALUE_	IMPNUM
1	2545235	1	2013	900030	2	982.8333	NA	1

[7] The dataset views in this document were generated using R, which represents missing values as "NA." Other statistical software packages may represent missing values as a period, blank space, or in another manner.

2	2545235	1	2013	900030	2	982.8333	NA	2
3	2545235	1	2013	900030	2	982.8333	NA	3
4	2545235	1	2013	900030	2	982.8333	NA	4
5	2545235	1	2013	900030	2	982.8333	NA	5
6	2545235	2	2013	900030	2	982.8333	NA	1
7	2545235	2	2013	900030	2	982.8333	NA	2
8	2545235	2	2013	900030	2	982.8333	NA	3
9	2545235	2	2013	900030	2	982.8333	NA	4
10	2545235	2	2013	900030	2	982.8333	NA	5

3.3.6 NTAXI-Federal and state tax information classified by tax unit

- There may exist multiple records per CU.
- There is one record per tax unit.
- Unique records are defined by the combination of NEWID and TAX UNIT.
- Variables include federal and state tax rates, tax liabilities, and various tax-related characteristics of the tax unit.

Note: Not all the variables within the NTAXI file are shown in the sample below.

	NEWID	TAX_UNIT	DEPCNT	FILESTAT	SRATE_CY	SRATE_PY
1	2545235	1	2	3	0.0000	0.0000
2	2545245	1	2	3	0.0000	0.0000
3	2545255	1	0	1	0.0000	0.0653
4	2545275	1	0	1	0.0575	0.0575
5	2545305	1	0	2	0.0475	0.0475
6	2545355	1	0	1	0.0758	0.0764
7	2545365	1	2	2	0.0930	0.0930
8	2545375	1	2	3	0.0000	0.0000
9	2545385	1	0	1	0.0610	0.0610
10	2545405	1	1	3	0.0400	0.0400

3.4 Interview annual files–EXPN

- Each EXPN file is different, with some having multiple records per CU and others with one record or no records per CU interviewed each quarter.
- Unique records are defined differently in each file.
- Common variables to all EXPN files: **NEWID** is the unique CU identifier. **QYEAR** defines the year and interview quarter for a given record. **SEQNO** is assigned sequentially during the interview as each expenditure record is recorded into the database. **ALCNO** is assigned sequentially for each record that has been allocated from one expenditure during the process. For example, a CU may report spending $50 on a

pair of men's pants and a shirt. The CE will allocate out that record into separate records, one for men's pants and shorts($30) and one for men's shirts($20). **REC ORIG** describes the origin of the record, i.e., whether the expenditure data was collected during the current interview, expenditure data collected for a current reference period in a previous interview, household inventory data collected during a previous interview that was pulled forward to the current, or data that was created by the CE processing edit system.

- Variables include those directly collected from the Interview survey questionnaire on expenditures and screener questions. Variables that are recalculated by CE are also included.

3.4.1 Three examples of EXPN files:

1 CLA-Clothing and sewing materials

Unique records in CLA are defined by the combination of NEWID, SEQNO, and ALCNO.

CLOTHYA is an item code variable for the type of clothing expenditure (suits, dresses, undergarments, etc.)

CLOTHMOA is an indicator variable for the month in which the purchase occurred.

CLOTHXA is the amount of the clothing expense.

Note: Not all the variables within the CLA file are shown below.

	QYEAR	NEWID	SEQNO	ALCNO	CLOTHYA	CLOTHMOA	CLOTHXA
1	20131	2511505	15	0	140	12	204
2	20131	2511555	13	0	270	11	100
3	20131	2511555	13	1	200	11	203
4	20131	2511555	13	2	210	11	40
5	20131	2511555	13	3	190	11	725
6	20131	2511565	15	0	275	12	100
7	20131	2511565	15	1	190	12	572
8	20131	2511565	15	2	150	12	78
9	20131	2511565	16	0	140	12	658
10	20131	2511565	17	0	200	12	759

2 IHB-Hospitalization and health insurance detailed questions

Unique records in IHB defined by NEWID, SEQNO, ALCNO, and HHIPDLIB.

HHIPDLIB is the policy number for the health insurance plan. This number is derived by the order in which the respondent lists their health insurance plans.

HHICOVQ is the number of people covered by health insurance plan.

QHI3MCX is the amount paid in health insurance premiums during the reference period (derived by CE).

HHICODE is an indicator variable for the type of insurance plan (HMO, FFS, Medicare supplement, or other special purpose plan).

Note: Not all the variables within the IHB file are shown below.

	QYEAR	NEWID	SEQNO	ALCNO	HHIPDLIB	HHICOVQ	QHI3MCX	HHICODE
1	20131	2511505	24	0	1	3	2	2
2	20131	2511505	25	0	2	2	2	2
3	20131	2511555	30	0	1	1	1075	1
4	20131	2511555	31	0	2	1	1297	1
5	20131	2511555	32	0	3	1	1297	1
6	20131	2511565	27	0	1	1	1	1
7	20131	2511565	29	0	3	1	1	4
8	20131	2511575	33	0	1	4	1	1
9	20131	2511585	35	0	1	2	1	2
10	20131	2511595	37	0	2	2	329	1

3 VEQ-Vehicle maintenance and repair

Unique records in VEQ defined by NEWID, SEQNO, and ALCNO.

VOPSERVY is an indicator variable that describes the type of maintenance or repair untaken.

VOPMOA is an indicator variable for the month in which the expense occurred.

VOPEXPX is the total cost of the maintenance or repair expense.

Note: Not all the variables within the VEQ file are shown below.

	QYEAR	NEWID	SEQNO	ALCNO	VOPSERVY	VOPMOA	VOPEXPX
1	20131	2511505	63	0	100	12	275
2	20131	2511575	61	0	100	11	103
3	20131	2511595	76	0	370	12	201
4	20131	2511625	39	0	115	11	819

5	20131	2511625	39	1	100	11	1
6	20131	2511625	39	2	110	11	1
7	20131	2511675	54	0	100	11	715
8	20131	2511695	126	0	100	12	808
9	20131	2511695	127	0	365	10	417
10	20131	2511695	128	0	365	11	382

4 Diary Survey Data

A Diary "quarter" refers to the calendar quarter in which the Diary booklet was placed in the home of the CU by the Census Field Representative. All Diary files are organized as quarterly files. Those include:

1. FMLD - a file with characteristics, income, weights, and summary level expenditures for the CU.
2. MEMD - a file with characteristics and income for each member in the CU.
3. EXPD - a detailed weekly expenditure file categorized by UCC.
4. DTBD - a detailed annual income file categorized by UCC.
5. DTID - a Consumer Unit imputed income file categorized by UCC.[8]

4.1 Diary files

- Most Diary files are analogous to Interview files.
- There are slightly different variable names in the Diary files.
- The summary level expenditures in the Diary FMLD file are not as extensive as those in the Interview FMLI file.

4.1.1 FMLD-CU characteristics, income, weights, and summary level expenditures

Analogous to FMLI file in Interview.

Note: Not all the variables within the FMLD file are shown below.

	NEWID	STRTMNTH	WEEKI	CUTENURE	AGE_REF	AGE2	FOODHOME	FINLWT21
1	1451201	4	1	1	70	50	52.43	34683.44
2	1451202	5	2	1	70	50	14.00	34683.44
3	1451211	4	1	4	33	1	0.00	31140.85
4	1451212	4	2	4	33	1	28.03	31140.85
5	1451221	4	1	3	88	65	0.00	40310.51

[8] Available beginning in 2004.

6	1451222	4	2	3	88	65	0.00	40310.51
7	1451241	5	1	1	43	19	152.65	35929.24
8	1451242	5	2	1	43	19	124.05	35929.24
9	1451261	4	1	2	31	1	108.78	33222.60
10	1451262	5	2	2	31	1	61.03	33222.60

4.1.2 MEMD-CU member characteristics and income

Analogous to MEMI file in Interview.

Note: Not all the variables within the MEMD file are shown below.

	NEWID	MEMBNO	AGE	EDUCA	MARITAL	EMPLTYPE	WAGEXM
1	1451201	1	70	5	1	NA	150
2	1451201	2	68	5	1	NA	150
3	1451202	1	70	5	1	NA	350
4	1451202	2	68	5	1	NA	160
5	1451211	1	33	5	3	NA	160
6	1451211	2	19	NA	5	NA	230
7	1451211	3	18	NA	5	NA	120
8	1451211	4	18	NA	5	NA	160
9	1451212	1	33	5	3	NA	270
10	1451212	2	19	NA	5	NA	190

4.1.3 EXPD-Detailed weekly expenditure file categorized by UCC

Analogous to MTBI files in Interview.

- There are multiple records per CU.
- There is one record per expenditure per purchase month/year.
- Unique records are not defined.
- The weekly amounts are recorded directly from listings in the completed Diary.

Note: Not all the variables within the EXPD file are shown below.

	NEWID	ALLOC	COST	GIFT	PUB_FLAG	EXPNSQDY	UCC
1	1451201	0	6.63	2	2	7	10210
2	1451201	0	1.40	2	2	1	20210
3	1451201	0	1.60	2	2	1	110210
4	1451201	0	42.80	2	2	1	170310
5	1451201	0	10.00	2	2	6	190111
6	1451201	0	24.00	2	2	5	190112
7	1451201	0	10.00	2	2	4	190211

8	1451201	0	2.00	2	2	5	190321
9	1451201	0	8.00	2	2	1	190322
10	1451201	0	2.00	2	2	3	190322

4.1.4 DTBD-Detailed annual income file categorized by UCC

Analogous to ITBI file in Interview.

Note: This DTBD example shows ALL variables that are in the full DTBD file.

	NEWID	UCC	AMOUNT	AMOUNT_	PUB_FLAG
1	1451201	900030	46918	NA	1
2	1451201	980000	46918	1	1
3	1451201	980010	2	NA	1
4	1451201	980020	70	NA	1
5	1451201	980040	2	NA	1
6	1451201	980060	2	NA	1
7	1451201	980070	46918	1	1
8	1451201	980090	100	NA	1
9	1451201	980210	100	NA	1
10	1451201	980230	100	NA	1

4.1.5 DTID-Consumer Unit imputed income file categorized by UCC.

Analogous to the ITII file.

Note: This DTID example shows ALL variables that are in the full DTID file.

	NEWID	UCC	PUB_FLAG	AMOUNT	AMOUNT_	IMPNUM
1	1451201	900030	1	46918	NA	1
2	1451201	900030	1	46918	NA	2
3	1451201	900030	1	46918	NA	3
4	1451201	900030	1	46918	NA	4
5	1451201	900030	1	46918	NA	5
6	1451201	980000	1	46918	1	1
7	1451201	980000	1	46918	1	2
8	1451201	980000	1	46918	1	3
9	1451201	980000	1	46918	1	4
10	1451201	980000	1	46918	1	5

5 Documentation

To aid in using the PUMD, the CE produces annual user guides and data dictionaries. Since the 2007 PUMD release, the user guides and data dictionaries, for both Interview and Diary data, have been

parsed out into separate documents. Prior to 2007, the user guides and data dictionaries were combined in one document. The user documentations include:

- A summary of changes from the previous release of CE PUMD.
- More in depth information of the PUMD files.
- Information on CE PUMD nondisclosure requirements.
- Formulas and procedures for estimating aggregate statistics from the data.
- Data reliability statements.
- And other information, including information on sampling design and cooperation levels.

The data dictionaries include a listing of all the CE PUMD variables and descriptions. User guides and data dictionaries dating back to the 1996 release, along with other miscellaneous documentations, can be accessed online at the CE website, http://www.bls.gov/cex/csxmicrodoc.htm.