



July 12, 2012 1

MEMORANDUM FOR Documentation

From: Franklin Silberstein  
Victimization and Expenditure Branch  
Demographic Statistical Methods Division  
U. S. Census Bureau



Subject: Sample Design of the 2012 Consumer Expenditures Diary (CED) Web  
Diary Study and SCIF Changes

## I. Introduction

This memorandum describes the CED Web Diary sample design, which is also described in POCCB change request #898.

In order to examine alternative means of collecting CED data, the Bureau of Labor Statistics plans to conduct a study of the Consumer Expenditures Diary done via the web. This study will allow a more complete analysis of the limitations and benefits of the Web Diary (Web Feasibility Test Statement of Work, October 2011). Allowing the respondent to enter data through the internet would certainly reduce collection cost. There is evidence of possible improvements, including accuracy of reporting (Couper, 2008), better response rates, and more flexibility. Once sample is selected and results come in, we will be better able to judge the possibility of using this method of data collection.

## II. Design for the CED Web Diary Sample

Due to time constraints, we will only use sample from the unit frame. We estimate needing at least 1,313 sample units in order to receive 500 completed total interviews. We will select reserve sample available in the unit frame to be in sample during October through December 2012. In selecting sample, we avoid selecting reserve sample already used in an earlier CED Feasibility Study (Census Bureau, 2007). Sample from that study is identified by sample date "888888". Later on in September 2009, eleven Stratification PSUs (SPSUs) were reduced from reserve sample (US Census, 2009). Therefore, before selecting sample for this study, it would be important to check and eliminate any of those SPSUs from possible sample.

Since the Census Bureau is undergoing a major Regional Office (RO) Management Restructuring, we want to only select sample in geographic areas that have finished realigning and have adequate time to adjust to the new RO Management Structure. As waves one through four of the RO Restructuring will have finished realigning by September, we included only geography with wave codes 1-4. As interviews start October 1, 2012 through December 31, 2012, we will only include CED placement day code (CEDPLCDAY) ranging from 274 to 365. We will identify any code as part of this web diary study with a sample date (SAMPDATE) "777777". The actual date of the interview will be identified by CEDPLCDAY, not SAMPDATE.

Reserve units are separate from regular CED production sample. However, both contain the same type of cases and both are nationally distributed. When designing the sample, the Census Bureau assigned enough reserve sample to accommodate potential research needs along with any need for sample expansions during the 2000 redesign (Census 2004). As the last feasibility test in 2006 also selected sample from the unit frame, this current test will do the same. The unit frame makes up about 80% of the CED reserve sample (as well as the regular CED sample) (U S Census Bureau, 2004). Unit-frame cases refer to CED cases obtained from the 2000 Census that typically have complete addresses that include a house number and street name.

DSD set up the following limitations on the sample. Concerned about the expense of providing field staff to set up and monitor each housing unit in the web diary study, they decided to only send out field staff where there were with at least five eligible units in a Field PSU. Hence, the study will only contain Field PSUs with at least five available units. In seeking a more representative sample of the nation, they decided against limiting sample to a single region or two. However, after further consideration, they decided to remove both Alaska and Hawaii units from the study after weighing the enormous cost of the field staff in those states versus the benefits of obtaining only slightly greater representation. Hence the study will only contain unit frame sample from all continental states and from Field PSUs where there are at least 5 eligible units, fully realigned in waves 1-4.

Since the prior CED Individual Feasibility Study conducted in 2006 used reserve sample designation E11, we decided to start using reserve sample from sample designations other than E11. Of the available sample under the above limitations, we found only 1,112 units in sample designation E12 and needed to seek remaining units from another sample designation.

Therefore, in need of at least 201 more units, we started looking at sample designation E13. In deciding sample from E13, we examined only E13 Field PSUs already selected for E12. Of those PSUs, we found 1088 available units. As we needed at least 201 units from 1088 available, we determined needing about 18.5% of available sample. To make that happen, we looked at 19 reduction codes from 101 available to select our sample. As  $101/19$  is approximately 5.315789, we wanted each code selected to be spaced approximately 5.315789 apart. Randomly picking a whole number starting digit between 0.00 and 5.315789, we looked at a table of random numbers between 000 and 532 and found 088, or 0.88. Rounding up to a whole number we arrived at a

starting number of 1. Adding multiples of 5.315789 to 1, we arrived at the following numbers: 1, 6.315789, 11.63158, 16.94737, 22.26316, 27.57895, 32.89474, 38.21053, 43.52632, 48.84211, 54.15789, 59.47368, 64.78947, 70.10526, 75.42105, 80.73684, 86.05263, 91.36842, 96.68421. After rounding to the nearest whole numbers, we arrived at the following 19 reduction group codes: '001', '006', '012', '017', '022', '028', '033', '038', '044', '049', '054', '059', '065', '070', '075', '081', '086', '091', and '097'. Using those 19 reduction group codes, we were able to find exactly 201 more units from Sample Designation E13.

Due to the web diary sample in E13 being represented by 19 reduction group codes, we will have to also adjust the base weights by a factor of 101/19 for every case. This will be done in a different memorandum.

### III. SCIF Changes

Three variables will be added to the SCIF: QTYPE, WDUSER, and WDPIN. The QTYPE code refers to the questionnaire type, will be one character in length, and DSMD will set that value to 1 for all values in the web diary, thus identifying all web diary sample. All other sample outside of the web diary will receive an empty value for QTYPE. The QTYPE code will be in recordtype RT2501 and located at position 146.

DSD will be responsible for populating WDUSER and WDPIN. The WDUSER code refers to the username of the web diary respondent. The WDPIN code refers to the password or pin code of the user. Used together, both WDUSER and WDPIN allow the respondent to gain access to the web diary. Both WDUSER and WDPIN codes will be in recordtype RT8500. WDUSER will be located in positions 50-55. The WDPIN also will be located in positions 61-66. For the Web Diary system test, DSD will provide DSMD with a complete test SCIF containing values for QTYPE, WDUSER, and WDPIN.

Regarding time frames, DSD has indicated that “the first production SCIF for the Web Diary Feasibility Test is scheduled to be delivered by DSMD to MCS on 8/28. DSD will populate RT8500 with WDUSER/WDPIN values, merge the SCIF, and send it back to DSMD for delivery to MCS by the 8/28 deadline.”

Research cases (QTYPE=1) should be assigned the following interview periods on the production SCIF:

| <u>Production Month</u> | <u>Research Interview Period</u> |
|-------------------------|----------------------------------|
| 201210                  | 201213                           |
| 201211                  | 201214                           |
| 201212                  | 201215                           |

References:

U. S. Census Bureau (2007), “Specification to Create Sample Control File for Consumer Expenditures Individual Diary Feasibility Field Test”, from Patrick Flanagan to Mackey Lewis, February 2006.

Bureau of Labor Statistics, “Web Diary Feasibility Test: Statement of Work”, October 13, 2011.

U. S. Census Bureau (2004). From Demographic Statistical Methods Division, “CE FACT SHEET (2000 Design),” November 10, 2004.

U. S. Census Bureau (2006). From Processing and Outputs Change Control Board, “Change Control #197”, approved February 16, 2006.

U. S. Census Bureau (2009). From Processing and Outputs Change Control Board, “Change Control #646”, approved September 31, 2009.

Miscellaneous:

This memorandum is stored in the VEB directory: “M:\ADC-SDE\VEB \CE\2012 Web Diary Test” with the name “CR898.docx”. As it also is an attachment to change request 898, it is also stored under the name “CR898.docx” with other change request attachments in the directory: “M:\SHARED\COMMON\POCCB\CR Attachment\”.

cc: P. Flanagan  
S. Barber  
S. Ash  
D. Castello  
R. Yoder  
J. Karanek