

## Supporting Statement

### B. Collection of Information Employing Statistical Methods

#### 1. Consumer Units

There are approximately 116.4 million consumer units (CUs) in the potential CE survey universe.<sup>1</sup> A CU is the unit from which we desire expenditure reports. It consists of all household members of a particular housing unit or other type of living quarters who are related by blood, marriage, adoption, or some other legal arrangement. The CU determination for unrelated persons is based on financial independence. The unrelated persons are considered separate CU(s) if they are responsible for paying their own expenses for two out of three of the following expense categories: shelter, food, and all other expenses. Approximately 97 percent of all occupied living quarters are a single CU.

For an overview of the CE sample design and the CU selection process, please refer to the 2007 CE Anthology article, "How Was My Household Selected for the Consumer Expenditure Survey? The Design and Selection of the Survey's Sample" by Sylvia Johnson-Herring and Susan King. This article is available upon request.

The following table shows estimated numbers of CUs for all 91 strata from which PSUs were selected.<sup>2</sup> (See the section below entitled "PSUs" for more information.)

Stratum Code	Estimated Number of CUs in Stratum	Stratum Code	Estimated Number of CUs in Stratum	Stratum Code	Estimated Number of CUs in Stratum	Stratum Code	Estimated Number of CUs in Stratum
A102	2,559,213	A313	1,055,780	A426	1,068,366	X224	1,082,708
A103	2,935,498	A316	2,181,835	A427	1,044,510	X226	1,137,025
A104	1,403,563	A318	1,950,037	A429	1,344,800	X228	946,721
A109	3,311,790	A319	1,737,397	A433	1,309,891	X232	1,392,266
A110	3,192,066	A320	1,603,061	X102	1,482,049	X336	1,287,497
A111	2,754,939	A321	1,313,016	X104	1,275,807	X338	1,221,154
A207	3,793,086	A419	5,113,753	X108	1,348,185	X340	1,398,206
A208	2,256,483	A420	1,657,499	X210	1,075,910	X342	1,550,352
A209	1,325,561	A422	2,911,099	X212	1,082,608	X344	1,281,584
A210	1,218,236	A423	1,470,056	X218	1,128,354	X346	1,276,391
A211	1,296,962	A424	1,163,649	X220	1,128,308	X350	1,114,492
A312	2,079,224	A425	1,212,368	X222	1,059,808	X352	1,247,393

<sup>11</sup> The number of CUs comes from dividing the Census Bureau's 2005 estimate of the number of people in the civilian non-institutional population (291 million) by the average number of people per CU (2.5).

<sup>22</sup> The number of CUs per stratum comes from allocating the nationwide total of 116.4 million CUs by each stratum's proportion of the nationwide population in the 2000 Census.

Stratum Code	Estimated Number of CUs in Stratum	Stratum Code	Estimated Number of CUs in Stratum
X354	1,193,368	Z322	943,804
X356	1,260,701	Z324	561,559
X358	1,238,172	Z426	271,401
X362	1,346,115	Z428	235,813
X364	1,367,570	Z430	325,837
X366	1,326,273	Z432	343,333
X368	1,219,167	<b>Total</b>	<b>116,400,000</b>
X472	1,090,016		
X474	1,081,970		
X482	1,145,328		
X484	1,164,280		
Y102	562,716		
Y104	654,250		
Y206	825,058		
Y208	981,226		
Y210	817,980		
Y212	980,953		
Y314	812,774		
Y316	938,987		
Y318	778,777		
Y320	819,967		
Y322	969,428		
Y324	776,087		
Y426	528,041		
Y428	482,697		
Y430	595,002		
Y432	575,530		
Z102	369,878		
Z104	575,351		
Z206	849,611		
Z208	624,732		
Z210	768,195		
Z212	912,352		
Z314	773,198		
Z316	733,707		
Z318	777,472		
Z320	1,020,768		

## Response Rates

The following table shows expected response rates for the Quarterly Interview and Diary Surveys based on 2005 response rates.

Category	Quarterly (quarter)	Diary (annual)
Total Sample Size	14,600	12,100
Total Type B and C Noninterviews (vacant, demolished, etc.):		
Number	2,701	2,238
Percent of Total Sample	18.5	18.5
Total Eligible Units	11,899	9,862
Total Type A Noninterviews		
Number	3,074	2,812
Percent of Total Eligible	25.8	28.5
Total Completed Interviews		
Number	8,825	7,050
Percent of Total Eligible	74.2	71.5

Note: For more information on the calculation of response rates, please see “Updating the Monthly Nonresponse Reports,” a document maintained in the Demographic Surveys Division of the Bureau of the Census.

## PSUs

The CE Survey uses a multistage stratified sample. The set of sample PSUs used for both the Quarterly Interview and Diary Surveys consists of 91 PSUs, 75 of which the BLS defined and selected for the CPI. The BLS derived these PSUs from a stratified design with the objective of selecting one PSU from each stratum. Stratification in the X, Y, and Z PSUs used a 5-variable geographic model whose independent variables are normalized longitude, the square of normalized longitude, normalized latitude, the square of normalized latitude, and percent urban consumer units.

The 91 CE PSUs fall within four categories:

<b>PSU Letter</b>	<b>Definitions</b>	<b>Number of PSUs</b>
A	Self-Representing Metropolitan CBSAs	21
X	Non-Self-Representing Metropolitan CBSAs	38
Y	Micropolitan CBSAs	16
Z	Non-CBSA PSUs	16

## 2. Sampling Within PSUs

To select living quarters within each PSU, the Census Bureau uses a four-frame approach accessing a combination of Census 2000 address lists, a specially constructed and continuously updated frame of new construction permits, and area segments for the sample selection. (The four frames are often called Unit, Area, GQ, and Permit.) The Area frame consists of units in blocks that are either in non-permit-issuing areas, or in areas for which the Census Bureau does not have reliable address list. Group quarters in Area frame blocks are automatically sampled along with the other units during Area frame listing operations. New construction in Area frame blocks is listed along with old construction; however if a unit is selected for sample from an Area frame block in a permit-issuing area, the interviewer asks what year the unit was built before beginning the interview. If the unit was built after 1999, then the unit is dropped from the Area frame sample because it has a chance of selection in the Permit frame. The Group Quarters (GQ) frame consists of GQ units in non-Area frame blocks which existed at the time of the 2000 Census. The Unit frame consists of non-GQ units in non-Area frame blocks which existed at the time of the 2000 Census. The Permit frame consists of permit new construction units in permit-issuing areas.

For more information on sampling within PSUs for the CE surveys, please refer to "Determining Within-PSU Sample Sizes for the Consumer Expenditure Survey" by David Swanson, Sharon Krieger, and Sylvia Johnson-Herring in the *Proceedings of the Section on Government Statistics*, American Statistical Association, 2003. This article is available upon request.

In general, since the sample sizes are small, we are using segments with clusters of size one to achieve maximum reliability of the estimates.

The Census Bureau uses an un-clustered systematic sample design to select the Unit frame sample from the 2000 census 100 percent detail file. The sample drawn from the detail file is augmented by sample from the new construction (Permit) frame. This frame consists of building permits which were authorized for housing units since the Census. A sample of newly constructed housing units represented by the building permits is drawn in each CE PSU.

The sample design for the unit frame stratifies and sorts HUs within an area with living quarters further subdivided into two categories: occupied HUs at the time of the census, and vacant HUs. The design provides for further stratification of occupied HUs by two 2000 census variables: a tenure-value-rent combination and household size. These variables were chosen following extensive research on within-PSU sampling prior to the sample redesign based on the 1990 Census. Because value and rent were only collected from a sample of housing units (the Long Form) in Census 2000, these two variables were imputed for those units not in the (Long Form) sample. Another change from the previous sample redesign is that rent and value categories were based on percentiles within counties, rather than pre-determined cutoffs. A single digit code was determined by tenure-value-rent. Then a second single digit code was determined based on the household size and the value of the first code. The concatenation of these two codes is the CE Unit frame stratum code.

Tables 1 and 2 provide the group definition (stratum code) of each variable and the associated group numbers for occupied HUs.

**Table 1: Tenure-Value-Rent Code**

<b>[Digit 1 of Unit Frame CE Stratification Code]</b>		
<b>Tenure</b>	<b>Value/Rent Percentile Range</b>	<b>Group #</b>
<b>Owner</b>	I (0 % - 24 %)	3
	II (25 % - 49 %)	4
	III (50 % - 74 %)	7
	IV (75 % - 99 %)	8
<b>Renter</b>	I (0 % - 24 %)	1
	II (25 % - 49 %)	2
	III (50 % - 74 %)	5
	IV (75 % - 99 %)	6

**Table 2: Household Size Code**  
**[Digit 2 of Unit Frame CE Stratification Code]**

<b>Group # for Tenure-Value-Rent Code</b>	<b>HU Size</b>	<b>Group # for HU Size</b>
1, 3, 5, or 7	1	1
	2	2
	3	3
	4+	4
2, 4, 6, or 8	1	4
	2	3
	3	2
	4+	1

The tenure-value-rent code and HU size code form the first and second digits, respectively, of the stratum code.

Table 3 below provides the strata definition and the corresponding strata codes for the vacant HUs:

**Table 3: Unit Frame CE Stratification Code for Vacant HUs**

Tenure	Value/Rent Percentile Range	Code for:	
		HU: Vacant for Sale/Rent	HU: Residual Vacant
Owner	I (0 % - 24 %)	30	99
	II (25 % - 49 %)	45	
	III (50 % - 74 %)	70	
	IV (75 % - 99 %)	85	
Renter	I (0 % - 24 %)	10	
	II (25 % - 49 %)	25	
	III (50 % - 74 %)	50	
	IV (75 % - 99 %)	65	

The Census Bureau forms Basic PSU Components (BPCs) from the intersections of the various demographic surveys’ stratification PSUs. (In the 2000 Sample Redesign, a BPC always consists of one or more counties.) The Census Bureau sorts the CE Unit frame in a given BPC by Urban/Rural classification, FIPS County code, CE Unit frame stratification code, Census Tract code, Census Tabulation Block code, Basic Street Address, and the Unit Sort Order<sup>3</sup> code. A systematic sample is then drawn from the sorted file.

The Area, GQ, and Permit frames are constructed from “measures”, where each measure is expected to contain four units. Measures are sorted geographically in the Area and GQ frames, and by Permit Date and Building Permit Office code in the Permit frame. Systematic samples of measures are selected independently in each frame. We expect the distribution of designated sample units among the four frames to be:

- Unit: 80%
- Area: 10%
- Permit: 9%
- GQ: 1%

Note, however, that the proportion of sample in the Area and Permit frames will increase as new construction is added to the sample.

Estimation

The estimation procedure for both the Diary and Quarterly Interview Surveys follow

<sup>3</sup> The Unit Sort Order code was devised by the Census Bureau in an attempt preserve geographic proximity between units in the sort.

well-known statistical principles in that the final weight for each sample CU is the product of the inverse of the probability of selection; a weight adjustment to account for noninterviews; and a regression-based adjustment that incorporates post-stratifications to adjust for population undercount at the same time controlling for extreme weights. For additional information on the sample design and estimation methodology used in the CE surveys, please refer to "Chapter 16, Consumer Expenditures and Income" in the *BLS Handbook of Methods*. This chapter is available upon request.

The BLS requires an average coefficient of variation of 10 percent or less covering 2 years of data for Expenditure Classification expenditures for each of the 39 market basket areas.

For additional information, please refer to Kenneth V. Dalton's memo to Chester E. Bowie, "Specifications for the Selection of CE/CPI Samples in PSUs Based on the 2000 Census:", June 28, 2002 and Alan R. Tupek's memo to Kenneth V. Dalton, "Calculations of Within-PSU Sampling Intervals for the Census 2000-Based Redesign of the Consumer Expenditure Surveys and the CPI Permit New Construction Housing Sample", November 11, 2002. These documents are available upon request.

3. In the CE Survey, keeping the noninterview rate at a low level requires special efforts, particularly from the Census Bureau Field staff. For each refusal case, the regional office sends a special letter to the address and assigns the case for follow-up by the program supervisor, supervisory field representative, or senior interviewer, taking into account time and cost considerations.

To adjust for those noninterviews that the field staff cannot convert to interviews, the sample design provides for a noninterview adjustment in the estimation procedure. The computer processing employs special techniques in the Quarterly Interview Survey to reference data provided in the previous interview, keeping recall problems and interview time to a minimum.

4. At the time of this submission, there are no plans to test any procedures or methods in FY2007.
5. The Census Bureau will collect the data. Within the Census Bureau, you may consult the following individuals and their area of expertise for further information.

Sample Design:	Karen King (301) 763-1974
Data Collection:	Howard McGowan (301) 763-5342