

MEMORANDUM FOR Kenneth V. Dalton
 Associate Commissioner
 Office of Prices and Living Conditions
 Bureau of Labor Statistics

From: Alan R. Tupek
 Chief, Demographic Statistical Methods Division
 Bureau of the Census

Subject: Calculation of Within-PSU Sampling Intervals for the
 Census 2000-Based Redesign of the Consumer Expenditure
 Surveys and the CPI Permit New Construction Housing Sample

I. Purpose of this Document

This document explains how the Census Bureau will calculate within-PSU sampling intervals for the Census 2000 -based redesign of the Consumer Expenditure Surveys (Quarterly Interview and Diary) and the CPI Permit New Construction Housing Sample. The calculations are based on instructions provided by the Bureau of Labor Statistics in reference [1].

II. Calculating Sampling Intervals for the Consumer Expenditure (CE) Surveys

There are four basic steps involved in calculating the sampling intervals for the CE surveys. Appendices 7-10 are the code for the four SAS programs written to accomplish these basic steps:

- Allocate the national target sample size of 7700 housing units (HUs) among the 102 stratification PSUs, attempting to make the allocation as close to proportional as possible, but subject to the constraints that each CPI Index Area have at least 80 HUs and the Z size (non-CBSA) PSUs have a total of 400 HUs.
- Calculate factors for inflating the target sample sizes to account for expected non-response. The factors will be based on CE response rates in the years 1999-2001.
- Calculate the PSU designated sample sizes (the PSU allocations inflated for non-response.)

- Calculate the within-PSU sampling interval for each PSU as the ratio of the PSU measure of size¹ to the PSU designated sample size.

A. Allocate the National Target Sample Size to the PSUs

1. Allocate the 7300 CBSA Housing Units (HUs) to the 36 CPI Areas

There are 36 CPI Areas. Each of the 28 self-representing A PSUs is its own CPI Area; and each of the eight region/size classes formed by the X and Y PSUs is a CPI Area. (The four regions are 1=Northeast, 2=Midwest, 3=South, and 4=West. The two size classes are X and Y. Thus the eight non-A CPI Areas are X100-X400 and Y100-Y400.)

We want the allocation of the 7300 HUs among the 36 areas to be as close as possible to population proportionality, but with the constraint that we must allocate each CPI Area a minimum of 80 HUs. We measure distance from proportionality as the sum of squared differences between each area's fraction of the total population across all strata and each area's allocated fraction of the total 7300 HUs. We want to minimize this sum.

Briefly, this least-squares minimization problem can be stated as:

$$\text{Minimize} \quad \sum_{i=1}^{36} \left(\frac{a_i}{7300} - \frac{p_i}{p} \right)^2$$

$$\text{Subject to} \quad a_i \geq 80, \quad i = 1, 2, \dots, 36$$

Where

a_i = units allocated to area i

p_i = population represented by area i

$$p = \sum_{i=1}^{36} p_i$$

We solve this problem using the SAS Procedure NLP, as suggested in reference [2]. The solution to the problem consists of the optimal values for the a_i .

See Appendix 1 for a listing of the CPI Area allocations.

¹ The measure of size is the projected number of housing units in 2005 (by county.) See Reference [3] for an explanation of the projection

2. Allocate to the X and Y PSUs Within Each Region

Once we have determined the CPI Area allocations, we sub-allocate within each X and Y CPI Area to the PSUs. Each PSU's sub-allocation is proportional to the fraction of the CPI Area total population represented by the PSU. Specifically, the ratio of the PSU sub-allocation to the CPI Area allocation is equal to the ratio of the population represented by the PSU to the CPI Area total population.

See Appendix 2 for a listing of the X and Y PSU allocations.

3. Allocate the 400 Non-CBSA HUs to the Z PSUs

We allocate the 400 HUs designated for the Z PSUs so that each Z PSU's allocation is proportional to the fraction of the total non-CBSA population represented by that PSU. Specifically, the ratio of the Z PSU allocation to 400 is equal to the ratio of the population represented by the Z PSU to the total non-CBSA population.

See the end of Appendix 2 for a listing of the Z PSU allocations.

B. Calculate the Non-participation Inflation Factors

In order to achieve the target of obtaining completed interviews from 7,700 housing units² (HUs,) we need to designate enough sample HUs to account for non-participants. We project the participation rates based on results from the CE Interview and Diary Surveys during the three calendar years 1999 – 2001.

The final inflation factors we use are determined at the CPI Area level, or at the region level for the Z PSUs. For brevity, within this section we use the term "PSU group" to refer to either type of grouping.

See Appendix 3 for a listing of the PSU group inflation factors.

Our procedure for calculating the non-participation inflation factors is as follows:

1. Group the 1990 design PSUs into PSU groups corresponding to the 2000 design CPI areas or region/size classes. Specifically:
 - a) Except for three in the Midwest region, each 1990 A PSU is also a 2000 A PSU, with the same PSU code and CPI Area code. Each of these is a PSU group by itself.

² Note that we expect to get more than 7,700 completed interviews, because some housing units (HUs) contain multiple consumer units (CUs.) We estimate a "CU inflation factor" of 1.05, so 7,700 HUs should yield 8,085 completed CU interviews ($7,700 \times 1.05 = 8,085$).

- b) The three Midwest 1990 A PSUs A212, A213, and A214 are X PSUs in the 2000 design, so these become part of the X200 PSU group.
 - c) All of the B, C, and D 1990 PSUs are grouped according to the first two characters in their PSU code. Then we convert B to X, C to Y, and D to Z. This results in eleven PSU groups: X100, X200, X300, X400, Y200, Y300, Y400, Z100, Z200, Z300, and Z400.
 - d) Notice that there are no 1990 PSUs which directly correspond to the 2000 CPI Area Y100. Therefore the inflation factor calculated for the X100 PSU group will also be applied to the Y100 CPI Area.
2. For each of the 39 PSU groups created in step 1, and for each of the two surveys (Interview and Diary,) calculate the overall participation rate in that PSU group during the period 1999 – 2001. The participation rate for the interview survey is the number of completed interviews divided by the number of attempted interviews. The participation rate for the Diary survey is the number of completed diaries divided by twice the number of possible participants (since each participant is supposed to complete two diaries.) Also calculate national participation rates for each of the two surveys during that period.
 3. In each PSU group, and for each survey, calculate a weighted average of the PSU group participation rate with the national participation rate:

$$AVGPSURATE = (0.25)(NATRATE) + (0.75)(PSURATE)$$

4. In each PSU group, find the minimum of the two survey average rates, and use the inverse of this number as the PSU group inflation factor. Also, in PSU groups where the CED participation rate is lower than the CEQ participation rate, calculate a CEQ sub-sampling take-every as the ratio of the CEQ rate to the CED rate. We will sub-sample the CEQ sample after the initial samples are selected, in order to reduce the CEQ workload in PSUs where we expect a better participation rate for CEQ than for CED. We are doing this only for CEQ (and not CED) because the cost of a CEQ interview is large compared with the cost of getting a completed Diary.

C. Calculate the PSU Designated Sample Sizes

The PSU designated sample size is TWICE the product of the PSU sub-allocation and the PSU inflation factor. We multiply by two because we need separate sample hits for CEQ and CED. We assign hits alternately to the two surveys' samples during within-PSU sampling.

See Appendix 4 for a list of the PSU designated sample sizes.

D. Calculate the PSU Sampling Intervals

1. Project 2005 Housing Unit Counts by County

We use the same modified projections of 2005 housing unit (HU) counts the Current Population Survey (CPS) and Survey of Income and Program Participation (SIPP) are using. Documentation of the projections may be found in reference [3]. We modified those projections for counties in North Dakota, West Virginia, and the District of Columbia. The North Dakota and West Virginia projected housing unit state totals were less than the Census 2000 housing unit counts for those two states. This did not seem reasonable, so we replaced the projections for those two states with the Census 2000 counts (at the county level.) For DC, the projection was deemed unrealistic, and we replaced it with an estimate of 268,504 housing units.

2. Summarize HU Counts to PSU Level

For each PSU in the CE sample, we sum the projected HU counts from the counties in that PSU. This sum is the projected PSU HU count.

3. Calculate PSU Sampling Intervals

The final PSU sampling interval for each PSU is the ratio of the projected PSU HU count to the PSU designated sample size calculated in C.

See Appendix 5 for a listing of the PSU sampling intervals.

III. Calculate the CPI Permit New Construction Housing Sample Sampling Intervals

A. Project Yearly Permit Activity in CPI Sample PSUs

We will project 2005 annual permit activity in the counties selected for the CPI Permit New Construction Housing Sample PSUs based on county-level counts from the permit files received in the Census Bureau Demographic Statistical Methods Division (DSMD) from the Census Bureau Manufacturing and Construction Division (MCD) each month (and additionally once a year for Building Permit Offices which report annually.) We will use the files from 1997 through 2001 (calendar years.) Projections will be done separately for each county, then summed over all counties in sample.

See Appendix 6 (not included for confidentiality reasons) for an explanation and listing of the 2005 permit count projections by PSU.

Appendix 11 is the SAS code for the program we use to calculate the projections and the national sampling interval.

B. Divide Projection by 1440 and Multiply by 4 to get Sampling Interval

The final sampling interval (which is the same for all PSUs) is the ratio of the 2005 projected number of permits in CPI sample areas to the desired annual sample size (1440), multiplied by the expected number of addresses per hit (4.) The national sampling interval is:

$$\frac{1,106,785}{1,440} \times 4 = 3,074.4028$$

C. Monitor Sample Size and Reduce When Necessary

DSMD will monitor the number of permits being selected for the CPI Permit New Construction Housing Sample each year, and reduce the sample if it gets significantly larger than 1,440 permits a year.

IV. Appendices

Appendix 1: Listing of Target Sample Size Allocations by CPI Area

Appendix 2: Listing of Target Sample Size Allocations by PSU

Appendix 3: Listing of CE Participation Rates and Calculated Inflation Factors by Region/Size Class

Appendix 4: Listing of PSU Designated Sample Sizes

Appendix 5: Listing of PSU Sampling Intervals

Appendix 6: Documentation of 2005 Permit Activity Projection for Counties in the CPI Permit New Construction Housing Sample (not included for confidentiality reasons)

Appendix 7: SAS Program to Allocate National Target Sample Size to PSUs

Appendix 8: SAS Program to Calculate PSU Inflation Factors

Appendix 9: SAS Program to Calculate PSU Designated Sample Sizes

Appendix 10: SAS Program to Calculate PSU Sampling Intervals

Appendix 11: SAS Program to Project 2005 Permit Counts by County and Calculate the National Sampling Interval for the CPI New Construction Housing Sample

V. References

[1] Memorandum to Bowie from Dalton, “Specifications for the Selection of CE/CPI Samples in PSUs Based on the 2000 Census,” June 28, 2002

[2] Johnson-Herring, et. al., “Determining Within-PSU Sample Sizes for the Consumer Expenditure Survey,” <draft>

[3] Memorandum for Documentation from Lawrence S. Cahoon, prepared by David Hall, “Updated County-Level Population and Housing Unit Projections (Doc. #3.2-?-?),” <draft>

VI. Contacts

If you have any questions about this memorandum, please contact one of the following:

Padraic Murphy

Phone: 301-763-2192

e-mail: Padraic.A.Murphy@census.gov

Stephen Ash

Phone: 301-763-4294

e-mail: Stephen.Eliot.Ash@census.gov

Karen King

Phone: 301-763-1974

e-mail: Karen.E.King@census.gov

Appendix 1: Listing of Target Sample Size Allocations by CPI Area

CE REDESIGN 2000
TARGET SAMPLE SIZE
ALLOCATIONS FOR CPI AREAS

| CPI_AREA | CPI_AREA_ ALLOCATION |
|----------|-------------------------|
| A102 | 168.78 |
| A103 | 194.62 |
| A104 | 80.00 |
| A109 | 220.45 |
| A110 | 212.23 |
| A111 | 182.22 |
| A207 | 253.50 |
| A208 | 147.99 |
| A209 | 80.00 |
| A210 | 80.00 |
| A211 | 82.11 |
| A312 | 135.82 |
| A313 | 80.00 |
| A316 | 142.87 |
| A318 | 126.95 |
| A319 | 112.35 |
| A320 | 103.13 |
| A321 | 80.00 |
| A419 | 344.18 |
| A420 | 106.86 |
| A422 | 192.94 |
| A423 | 93.99 |
| A424 | 80.00 |
| A425 | 80.00 |
| A426 | 80.00 |
| A427 | 80.00 |
| A429 | 85.39 |
| A433 | 80.00 |
| X100 | 302.33 |
| X200 | 696.78 |
| X300 | 1342.32 |
| X400 | 445.80 |
| Y100 | 80.00 |
| Y200 | 240.60 |
| Y300 | 342.96 |
| Y400 | 142.83 |
| | ===== |
| | 7300.00 |

Appendix 2: Listing of Target Sample Size Allocations by PSU

CE REDESIGN 2000
TARGET SAMPLE SIZE
ALLOCATIONS FOR X- AND Y-SIZE PSUS

CPI_AREA=X100

| BLSPSU2K | PSU_ALLOCATION |
|----------|----------------|
| X102 | 99.477 |
| X104 | 70.190 |
| X106 | 57.713 |
| X108 | 74.950 |
| ----- | ----- |
| CPI_AREA | 302.329 |

CPI_AREA=X200

| BLSPSU2K | PSU_ALLOCATION |
|----------|----------------|
| X210 | 60.838 |
| X212 | 67.288 |
| X214 | 81.062 |
| X216 | 32.641 |
| X218 | 76.711 |
| X220 | 66.968 |
| X222 | 51.719 |
| X224 | 48.761 |
| X226 | 69.047 |
| X228 | 53.394 |
| X230 | 38.940 |
| X232 | 49.408 |
| ----- | ----- |
| CPI_AREA | 696.776 |

CPI_AREA=X300

| BLSPSU2K | PSU_ALLOCATION |
|----------|----------------|
| X334 | 76.9139 |
| X336 | 79.9226 |
| X338 | 79.3661 |
| X340 | 81.7532 |
| X342 | 74.3940 |
| X344 | 83.2174 |
| X346 | 75.3741 |
| X348 | 54.9928 |

Appendix 2: Listing of Target Sample Size Allocations by PSU

CE REDESIGN 2000
TARGET SAMPLE SIZE
ALLOCATIONS FOR X- AND Y-SIZE PSUS

CPI_AREA=X300
(continued)

| BLSPSU2K | PSU_ALLOCATION |
|----------|----------------|
| X350 | 63.92 |
| X352 | 81.63 |
| X354 | 82.28 |
| X356 | 82.98 |
| X358 | 81.48 |
| X360 | 61.34 |
| X362 | 74.39 |
| X364 | 83.00 |
| X366 | 42.09 |
| X368 | 83.28 |
| ----- | ----- |
| CPI_AREA | 1342.32 |

CPI_AREA=X400

| BLSPSU2K | PSU_ALLOCATION |
|----------|----------------|
| X470 | 72.832 |
| X472 | 47.220 |
| X474 | 63.038 |
| X476 | 47.159 |
| X478 | 69.271 |
| X480 | 49.186 |
| X482 | 48.631 |
| X484 | 48.466 |
| ----- | ----- |
| CPI_AREA | 445.801 |

CPI_AREA=Y100

| BLSPSU2K | PSU_ALLOCATION |
|----------|----------------|
| Y102 | 36.9914 |
| Y104 | 43.0086 |
| ----- | ----- |
| CPI_AREA | 80.0000 |

Appendix 2: Listing of Target Sample Size Allocations by PSU

CE REDESIGN 2000
 TARGET SAMPLE SIZE
 ALLOCATIONS FOR X- AND Y-SIZE PSUS

CPI_AREA=Y200

| BLSPSU2K | PSU_ALLOCATION |
|----------|----------------|
| Y206 | 55.062 |
| Y208 | 65.484 |
| Y210 | 54.589 |
| Y212 | 65.465 |
| ----- | ----- |
| CPI_AREA | 240.600 |

CPI_AREA=Y300

| BLSPSU2K | PSU_ALLOCATION |
|----------|----------------|
| Y314 | 54.700 |
| Y316 | 63.194 |
| Y318 | 52.412 |
| Y320 | 55.184 |
| Y322 | 65.243 |
| Y324 | 52.231 |
| ----- | ----- |
| CPI_AREA | 342.963 |

CPI_AREA=Y400

| BLSPSU2K | PSU_ALLOCATION |
|----------|----------------|
| Y426 | 34.58 |
| Y428 | 31.61 |
| Y430 | 38.96 |
| Y432 | 37.69 |
| ----- | ----- |
| CPI_AREA | 142.83 |
| | ===== |
| | 3593.62 |

Appendix 2: Listing of Target Sample Size Allocations by PSU

CE REDESIGN 2000
TARGET SAMPLE SIZE
ALLOCATIONS FOR Z-SIZE PSUS

| BLSPSU2K | PSU_ALLOCATION |
|----------|----------------|
| Z102 | 14.701 |
| Z104 | 22.106 |
| Z206 | 33.625 |
| Z208 | 24.830 |
| Z210 | 30.532 |
| Z212 | 36.261 |
| Z314 | 30.730 |
| Z316 | 29.161 |
| Z318 | 30.900 |
| Z320 | 40.570 |
| Z322 | 37.511 |
| Z324 | 22.319 |
| Z426 | 10.787 |
| Z428 | 9.372 |
| Z430 | 12.950 |
| Z432 | 13.646 |
| | ===== |
| | 400.000 |

Appendix 3: Listing of CE Response Rates and Calculated Inflation Factors by Region/Size Class

CE REDESIGN 2000
PARTICIPATION RATES AND INFLATION FACTORS
BY REGION/SIZE CLASS

| PSU GROUP | CEQ PARTICIPATION RATE | CEQ NATIONAL PARTICIPATION RATE | CEQ WEIGHTED AVERAGE RATE | CED PARTICIPATION RATE | CED NATIONAL PARTICIPATION RATE | CED WEIGHTED AVERAGE RATE | INFLATION FACTOR USED | CEQ SUMSAMPLING TAKE-EVERY |
|-----------|------------------------|---------------------------------|---------------------------|------------------------|---------------------------------|---------------------------|-----------------------|----------------------------|
| A102 | 0.55420 | 0.64988 | 0.57812 | 0.49284 | 0.62024 | 0.52469 | 1.90590 | 1.10183 |
| A103 | 0.69146 | 0.64988 | 0.68106 | 0.72693 | 0.62024 | 0.70026 | 1.46829 | 1.00000 |
| A104 | 0.66183 | 0.64988 | 0.65884 | 0.66287 | 0.62024 | 0.65221 | 1.53324 | 1.01017 |
| A109 | 0.60870 | 0.64988 | 0.61899 | 0.49392 | 0.62024 | 0.52550 | 1.90296 | 1.17791 |
| A110 | 0.66438 | 0.64988 | 0.66076 | 0.63752 | 0.62024 | 0.63320 | 1.57929 | 1.04353 |
| A111 | 0.65113 | 0.64988 | 0.65082 | 0.62884 | 0.62024 | 0.62669 | 1.59568 | 1.03850 |
| A207 | 0.60519 | 0.64988 | 0.61637 | 0.50191 | 0.62024 | 0.53149 | 1.88151 | 1.15970 |
| A208 | 0.65917 | 0.64988 | 0.65684 | 0.67039 | 0.62024 | 0.65785 | 1.52243 | 1.00000 |
| A209 | 0.64473 | 0.64988 | 0.64602 | 0.66603 | 0.62024 | 0.65458 | 1.54794 | 1.00000 |
| A210 | 0.68197 | 0.64988 | 0.67395 | 0.62427 | 0.62024 | 0.62326 | 1.60447 | 1.08133 |
| A211 | 0.71021 | 0.64988 | 0.69513 | 0.84375 | 0.62024 | 0.78787 | 1.43858 | 1.00000 |
| A312 | 0.65031 | 0.64988 | 0.65020 | 0.61212 | 0.62024 | 0.61415 | 1.62828 | 1.05870 |
| A313 | 0.65504 | 0.64988 | 0.65375 | 0.51782 | 0.62024 | 0.54343 | 1.84018 | 1.20301 |
| A316 | 0.67569 | 0.64988 | 0.66924 | 0.63666 | 0.62024 | 0.63255 | 1.58089 | 1.05800 |
| A318 | 0.68746 | 0.64988 | 0.67806 | 0.65410 | 0.62024 | 0.64564 | 1.54886 | 1.05022 |
| A319 | 0.71374 | 0.64988 | 0.69778 | 0.68305 | 0.62024 | 0.66735 | 1.49847 | 1.04560 |
| A320 | 0.63616 | 0.64988 | 0.63959 | 0.58473 | 0.62024 | 0.59361 | 1.68461 | 1.07746 |
| A321 | 0.68176 | 0.64988 | 0.67379 | 0.61142 | 0.62024 | 0.61362 | 1.62967 | 1.09805 |
| A419 | 0.65677 | 0.64988 | 0.65505 | 0.63379 | 0.62024 | 0.63040 | 1.58629 | 1.03910 |
| A420 | 0.58660 | 0.64988 | 0.60242 | 0.57124 | 0.62024 | 0.58349 | 1.71382 | 1.03244 |
| A422 | 0.69853 | 0.64988 | 0.68636 | 0.73390 | 0.62024 | 0.70548 | 1.45696 | 1.00000 |
| A423 | 0.68044 | 0.64988 | 0.67280 | 0.75000 | 0.62024 | 0.71756 | 1.48633 | 1.00000 |
| A424 | 0.56673 | 0.64988 | 0.58751 | 0.63311 | 0.62024 | 0.62989 | 1.70209 | 1.00000 |
| A425 | 0.75083 | 0.64988 | 0.72559 | 0.75054 | 0.62024 | 0.71797 | 1.39282 | 1.01062 |
| A426 | 0.64114 | 0.64988 | 0.64333 | 0.58245 | 0.62024 | 0.59190 | 1.68948 | 1.08689 |
| A427 | 0.67037 | 0.64988 | 0.66524 | 0.64931 | 0.62024 | 0.64204 | 1.55753 | 1.03613 |
| A429 | 0.61357 | 0.64988 | 0.62264 | 0.60151 | 0.62024 | 0.60619 | 1.64964 | 1.02714 |
| A433 | 0.65863 | 0.64988 | 0.65644 | 0.69811 | 0.62024 | 0.67864 | 1.52337 | 1.00000 |
| X100 | 0.68091 | 0.64988 | 0.67315 | 0.68361 | 0.62024 | 0.66777 | 1.49753 | 1.00807 |
| X200 | 0.70173 | 0.64988 | 0.68876 | 0.65887 | 0.62024 | 0.64921 | 1.54033 | 1.06092 |
| X300 | 0.64185 | 0.64988 | 0.64385 | 0.59355 | 0.62024 | 0.60022 | 1.66605 | 1.07269 |
| X400 | 0.64113 | 0.64988 | 0.64331 | 0.65653 | 0.62024 | 0.64746 | 1.55445 | 1.00000 |
| Y100 | 0.68091 | 0.64988 | 0.67315 | 0.68361 | 0.62024 | 0.66777 | 1.49753 | 1.00807 |
| Y200 | 0.69125 | 0.64988 | 0.68090 | 0.67328 | 0.62024 | 0.66002 | 1.51510 | 1.03164 |
| Y300 | 0.61565 | 0.64988 | 0.62421 | 0.55663 | 0.62024 | 0.57253 | 1.74663 | 1.09026 |
| Y400 | 0.66063 | 0.64988 | 0.65794 | 0.68908 | 0.62024 | 0.67187 | 1.51988 | 1.00000 |
| Z100 | 0.58890 | 0.64988 | 0.60414 | 0.63722 | 0.62024 | 0.63297 | 1.65523 | 1.00000 |
| Z200 | 0.53256 | 0.64988 | 0.56189 | 0.49407 | 0.62024 | 0.52561 | 1.90255 | 1.06902 |
| Z300 | 0.60390 | 0.64988 | 0.61540 | 0.52217 | 0.62024 | 0.54669 | 1.82919 | 1.12567 |
| Z400 | 0.58273 | 0.64988 | 0.59952 | 0.59038 | 0.62024 | 0.59784 | 1.67269 | 1.00281 |

Notes on Calculations:

For each survey,

$$[\text{Weighted Average Rate}] = (0.75)[\text{PSU Rate}] + (0.25)[\text{National Rate}]$$

Then,

$$[\text{Inflation Factor Used}] = 1 / \text{minimum}\{ [\text{CEQ Avg Rate}], [\text{CED Avg Rate}] \}$$

Appendix 4: Listing of PSU Designated Sample Sizes

CE 2000 REDESIGN DESIGNATED SAMPLE SIZES

| STRAT PSU | ALLOCATED TARGET SAMPLE SIZE | NON-RESPONSE INFLATION FACTOR | DESIGNATED SAMPLE SIZE |
|--------------|---------------------------------------|-------------------------------------|------------------------------|
| A102 | 168.778 | 1.90590 | 643.35 |
| A103 | 194.615 | 1.46829 | 571.50 |
| A104 | 80.000 | 1.53324 | 245.32 |
| A109 | 220.452 | 1.90296 | 839.02 |
| A110 | 212.232 | 1.57929 | 670.35 |
| A111 | 182.217 | 1.59568 | 581.52 |
| A207 | 253.500 | 1.88151 | 953.92 |
| A208 | 147.992 | 1.52243 | 450.62 |
| A209 | 80.000 | 1.54794 | 247.67 |
| A210 | 80.000 | 1.60447 | 256.72 |
| A211 | 82.109 | 1.43858 | 236.24 |
| A312 | 135.821 | 1.62828 | 442.31 |
| A313 | 80.000 | 1.84018 | 294.43 |
| A316 | 142.866 | 1.58089 | 451.71 |
| A318 | 126.951 | 1.54886 | 393.26 |
| A319 | 112.350 | 1.49847 | 336.71 |
| A320 | 103.126 | 1.68461 | 347.46 |
| A321 | 80.000 | 1.62967 | 260.75 |
| A419 | 344.180 | 1.58629 | 1091.94 |
| A420 | 106.864 | 1.71382 | 366.29 |
| A422 | 192.940 | 1.45696 | 562.21 |
| A423 | 93.994 | 1.48633 | 279.41 |
| A424 | 80.000 | 1.70209 | 272.33 |
| A425 | 80.000 | 1.39282 | 222.85 |
| A426 | 80.000 | 1.68948 | 270.32 |
| A427 | 80.000 | 1.55753 | 249.20 |
| A429 | 85.393 | 1.64964 | 281.74 |
| A433 | 80.000 | 1.52337 | 243.74 |
| X102 | 99.477 | 1.49753 | 297.94 |
| X104 | 70.190 | 1.49753 | 210.22 |
| X106 | 57.713 | 1.49753 | 172.85 |
| X108 | 74.950 | 1.49753 | 224.48 |
| X210 | 60.838 | 1.54033 | 187.42 |
| X212 | 67.288 | 1.54033 | 207.29 |
| X214 | 81.062 | 1.54033 | 249.72 |
| X216 | 32.641 | 1.54033 | 100.56 |
| X218 | 76.711 | 1.54033 | 236.32 |
| X220 | 66.968 | 1.54033 | 206.31 |
| X222 | 51.719 | 1.54033 | 159.33 |
| X224 | 48.761 | 1.54033 | 150.22 |

Note on Calculation:

Designated Sample Size = 2 x [Allocated Target Sample Size] x [Non-response Inflation Factor]
(Factor of 2 accounts for both Interview and Diary surveys needing the same sample size)

Appendix 4: Listing of PSU Designated Sample Sizes

| | | | |
|------|--------|---------|--------|
| X226 | 69.047 | 1.54033 | 212.71 |
|------|--------|---------|--------|

Note on Calculation:

Designated Sample Size = 2 x [Allocated Target Sample Size] x [Non-response Inflation Factor]
(Factor of 2 accounts for both Interview and Diary surveys needing the same sample size)

Appendix 4: Listing of PSU Designated Sample Sizes

CE 2000 REDESIGN DESIGNATED SAMPLE SIZES

| STRAT PSU | ALLOCATED TARGET SAMPLE SIZE | NON-RESPONSE INFLATION FACTOR | DESIGNATED SAMPLE SIZE |
|--------------|---------------------------------------|-------------------------------------|------------------------------|
| X228 | 53.3938 | 1.54033 | 164.488 |
| X230 | 38.9397 | 1.54033 | 119.960 |
| X232 | 49.4079 | 1.54033 | 152.209 |
| X334 | 76.9139 | 1.66605 | 256.284 |
| X336 | 79.9226 | 1.66605 | 266.310 |
| X338 | 79.3661 | 1.66605 | 264.455 |
| X340 | 81.7532 | 1.66605 | 272.409 |
| X342 | 74.3940 | 1.66605 | 247.888 |
| X344 | 83.2174 | 1.66605 | 277.288 |
| X346 | 75.3741 | 1.66605 | 251.154 |
| X348 | 54.9928 | 1.66605 | 183.241 |
| X350 | 63.9224 | 1.66605 | 212.995 |
| X352 | 81.6296 | 1.66605 | 271.998 |
| X354 | 82.2828 | 1.66605 | 274.174 |
| X356 | 82.9765 | 1.66605 | 276.486 |
| X358 | 81.4750 | 1.66605 | 271.483 |
| X360 | 61.3410 | 1.66605 | 204.394 |
| X362 | 74.3917 | 1.66605 | 247.880 |
| X364 | 82.9983 | 1.66605 | 276.558 |
| X366 | 42.0897 | 1.66605 | 140.247 |
| X368 | 83.2806 | 1.66605 | 277.499 |
| X470 | 72.8318 | 1.55445 | 226.427 |
| X472 | 47.2201 | 1.55445 | 146.803 |
| X474 | 63.0378 | 1.55445 | 195.979 |
| X476 | 47.1591 | 1.55445 | 146.613 |
| X478 | 69.2706 | 1.55445 | 215.356 |
| X480 | 49.1856 | 1.55445 | 152.913 |
| X482 | 48.6307 | 1.55445 | 151.188 |
| X484 | 48.4659 | 1.55445 | 150.676 |
| Y102 | 36.9914 | 1.49753 | 110.791 |
| Y104 | 43.0086 | 1.49753 | 128.813 |
| Y206 | 55.0616 | 1.51510 | 166.848 |
| Y208 | 65.4837 | 1.51510 | 198.429 |
| Y210 | 54.5892 | 1.51510 | 165.416 |
| Y212 | 65.4655 | 1.51510 | 198.374 |
| Y314 | 54.6998 | 1.74663 | 191.081 |
| Y316 | 63.1940 | 1.74663 | 220.753 |
| Y318 | 52.4118 | 1.74663 | 183.088 |
| Y320 | 55.1839 | 1.74663 | 192.772 |
| Y322 | 65.2426 | 1.74663 | 227.910 |

Note on Calculation:

Designated Sample Size = 2 x [Allocated Target Sample Size] x [Non-response Inflation Factor]
(Factor of 2 accounts for both Interview and Diary surveys needing the same sample size)

Appendix 4: Listing of PSU Designated Sample Sizes

| | | | |
|------|---------|---------|---------|
| Y324 | 52.2307 | 1.74663 | 182.456 |
|------|---------|---------|---------|

Note on Calculation:

Designated Sample Size = 2 x [Allocated Target Sample Size] x [Non-response Inflation Factor]
(Factor of 2 accounts for both Interview and Diary surveys needing the same sample size)

Appendix 4: Listing of PSU Designated Sample Sizes

CE 2000 REDESIGN
DESIGNATED SAMPLE SIZES

| STRAT PSU | ALLOCATED TARGET SAMPLE SIZE | NON-RESPONSE INFLATION FACTOR | DESIGNATED SAMPLE SIZE |
|--------------|---------------------------------------|-------------------------------------|------------------------------|
| Y426 | 34.5756 | 1.51988 | 105.10 |
| Y428 | 31.6066 | 1.51988 | 96.08 |
| Y430 | 38.9602 | 1.51988 | 118.43 |
| Y432 | 37.6852 | 1.51988 | 114.55 |
| Z102 | 14.7006 | 1.65523 | 48.67 |
| Z104 | 22.1060 | 1.65523 | 73.18 |
| Z206 | 33.6248 | 1.90255 | 127.95 |
| Z208 | 24.8297 | 1.90255 | 94.48 |
| Z210 | 30.5316 | 1.90255 | 116.18 |
| Z212 | 36.2610 | 1.90255 | 137.98 |
| Z314 | 30.7304 | 1.82919 | 112.42 |
| Z316 | 29.1609 | 1.82919 | 106.68 |
| Z318 | 30.9002 | 1.82919 | 113.04 |
| Z320 | 40.5699 | 1.82919 | 148.42 |
| Z322 | 37.5111 | 1.82919 | 137.23 |
| Z324 | 22.3189 | 1.82919 | 81.65 |
| Z426 | 10.7867 | 1.67269 | 36.09 |
| Z428 | 9.3723 | 1.67269 | 31.35 |
| Z430 | 12.9503 | 1.67269 | 43.32 |
| Z432 | 13.6456 | 1.67269 | 45.65 |
| | | | ===== |
| | | | 25028.79 |

Note on Calculation:

Designated Sample Size = 2 x [Allocated Target Sample Size] x [Non-response Inflation Factor]
(Factor of 2 accounts for both Interview and Diary surveys needing the same sample size)

Appendix 5: Listing of PSU Sampling Intervals

CE 2000 REDESIGN WITHIN-PSU SAMPLING INTERVALS

| STRAT PSU | PROJECTED 2005 HU COUNT | DESIGNATED SAMPLE SIZE | PSU SAMPLING INTERVAL |
|--------------|-------------------------------|------------------------------|--------------------------|
| A102 | 2644191 | 643.35 | 4,110.0307 |
| A103 | 3011714 | 571.50 | 5,269.8147 |
| A104 | 1104734 | 245.32 | 4,503.2623 |
| A109 | 3394801 | 839.02 | 4,046.1342 |
| A110 | 3102041 | 670.35 | 4,627.4896 |
| A111 | 2735086 | 581.52 | 4,703.3136 |
| A207 | 3656389 | 953.92 | 3,833.0045 |
| A208 | 2261974 | 450.62 | 5,019.7460 |
| A209 | 1182324 | 247.67 | 4,773.7656 |
| A210 | 1264006 | 256.72 | 4,923.7661 |
| A211 | 1315627 | 236.24 | 5,569.0242 |
| A312 | 2099845 | 442.31 | 4,747.4739 |
| A313 | 1116457 | 294.43 | 3,791.9436 |
| A316 | 2300308 | 451.71 | 5,092.4152 |
| A318 | 2016412 | 393.26 | 5,127.4678 |
| A319 | 1834109 | 336.71 | 5,447.2015 |
| A320 | 1811318 | 347.46 | 5,213.0894 |
| A321 | 1300498 | 260.75 | 4,987.5927 |
| A419 | 4712837 | 1091.94 | 4,316.0205 |
| A420 | 1597944 | 366.29 | 4,362.4799 |
| A422 | 2946667 | 562.21 | 5,241.2260 |
| A423 | 1585272 | 279.41 | 5,673.6122 |
| A424 | 1156037 | 272.33 | 4,244.9291 |
| A425 | 995210 | 222.85 | 4,465.7914 |
| A426 | 329978 | 270.32 | 1,220.7088 |
| A427 | 134075 | 249.20 | 538.0122 |
| A429 | 1553094 | 281.74 | 5,512.5904 |
| A433 | 1213534 | 243.74 | 4,978.8245 |
| X102 | 153879 | 297.94 | 516.4790 |
| X104 | 294956 | 210.22 | 1,403.0638 |
| X106 | 437022 | 172.85 | 2,528.2766 |
| X108 | 51073 | 224.48 | 227.5182 |
| X210 | 211424 | 187.42 | 1,128.0763 |
| X212 | 721424 | 207.29 | 3,480.2297 |
| X214 | 111463 | 249.72 | 446.3440 |
| X216 | 69517 | 100.56 | 691.3239 |
| X218 | 837510 | 236.32 | 3,543.9343 |
| X220 | 58658 | 206.31 | 284.3260 |
| X222 | 114574 | 159.33 | 719.1100 |
| X224 | 97159 | 150.22 | 646.7919 |
| X226 | 770985 | 212.71 | 3,624.5938 |
| X228 | 87590 | 164.49 | 532.4998 |

Note on Calculation:

Sampling Interval = [Projected HU Count] / [Designated Sample Size]

Appendix 5: Listing of PSU Sampling Intervals

CE 2000 REDESIGN WITHIN-PSU SAMPLING INTERVALS

| STRAT PSU | PROJECTED 2005 HU COUNT | DESIGNATED SAMPLE SIZE | PSU SAMPLING INTERVAL |
|--------------|-------------------------------|------------------------------|--------------------------|
| X230 | 167478 | 119.960 | 1,396.1149 |
| X232 | 209848 | 152.209 | 1,378.6836 |
| X334 | 368870 | 256.284 | 1,439.3000 |
| X336 | 139926 | 266.310 | 525.4255 |
| X338 | 726725 | 264.455 | 2,748.0055 |
| X340 | 292699 | 272.409 | 1,074.4816 |
| X342 | 451702 | 247.888 | 1,822.2038 |
| X344 | 176654 | 277.288 | 637.0773 |
| X346 | 677261 | 251.154 | 2,696.5993 |
| X348 | 97410 | 183.241 | 531.5946 |
| X350 | 139446 | 212.995 | 654.6899 |
| X352 | 278981 | 271.998 | 1,025.6738 |
| X354 | 568413 | 274.174 | 2,073.1811 |
| X356 | 87650 | 276.486 | 317.0145 |
| X358 | 299860 | 271.483 | 1,104.5279 |
| X360 | 82142 | 204.394 | 401.8809 |
| X362 | 136875 | 247.880 | 552.1818 |
| X364 | 540039 | 276.558 | 1,952.7130 |
| X366 | 251484 | 140.247 | 1,793.1529 |
| X368 | 437673 | 277.499 | 1,577.2055 |
| X470 | 817875 | 226.427 | 3,612.0868 |
| X472 | 138971 | 146.803 | 946.6497 |
| X474 | 703800 | 195.979 | 3,591.2093 |
| X476 | 79841 | 146.613 | 544.5692 |
| X478 | 300935 | 215.356 | 1,397.3849 |
| X480 | 75991 | 152.913 | 496.9545 |
| X482 | 114398 | 151.188 | 756.6586 |
| X484 | 86486 | 150.676 | 573.9863 |
| Y102 | 59102 | 110.791 | 533.4531 |
| Y104 | 50702 | 128.813 | 393.6088 |
| Y206 | 45581 | 166.848 | 273.1889 |
| Y208 | 21881 | 198.429 | 110.2712 |
| Y210 | 11881 | 165.416 | 71.8248 |
| Y212 | 13924 | 198.374 | 70.1907 |
| Y314 | 54340 | 191.081 | 284.3822 |
| Y316 | 20131 | 220.753 | 91.1924 |
| Y318 | 16683 | 183.088 | 91.1201 |
| Y320 | 21601 | 192.772 | 112.0547 |
| Y322 | 47787 | 227.910 | 209.6752 |
| Y324 | 19898 | 182.456 | 109.0566 |
| Y426 | 28963 | 105.102 | 275.5706 |
| Y428 | 58793 | 96.077 | 611.9384 |

Note on Calculation:

Sampling Interval = [Projected HU Count] / [Designated Sample Size]

Appendix 5: Listing of PSU Sampling Intervals

CE 2000 REDESIGN WITHIN-PSU SAMPLING INTERVALS

| STRAT PSU | PROJECTED 2005 HU COUNT | DESIGNATED SAMPLE SIZE | PSU SAMPLING INTERVAL |
|--------------|-------------------------------|------------------------------|--------------------------|
| Y430 | 48781 | 118.430 | 411.8970 |
| Y432 | 95340 | 114.554 | 832.2689 |
| Z102 | 29593 | 48.666 | 608.0841 |
| Z104 | 43955 | 73.181 | 600.6324 |
| Z206 | 21856 | 127.946 | 170.8223 |
| Z208 | 18482 | 94.479 | 195.6195 |
| Z210 | 12456 | 116.176 | 107.2170 |
| Z212 | 43005 | 137.977 | 311.6832 |
| Z314 | 12140 | 112.423 | 107.9847 |
| Z316 | 31685 | 106.681 | 297.0057 |
| Z318 | 13167 | 113.045 | 116.4759 |
| Z320 | 37531 | 148.420 | 252.8699 |
| Z322 | 62302 | 137.230 | 453.9979 |
| Z324 | 11050 | 81.651 | 135.3321 |
| Z426 | 31493 | 36.086 | 872.7299 |
| Z428 | 9868 | 31.354 | 314.7304 |
| Z430 | 13671 | 43.323 | 315.5567 |
| Z432 | 8382 | 45.650 | 183.6157 |

Note on Calculation:

Sampling Interval = [Projected HU Count] / [Designated Sample Size]

Appendix 7: SAS Program to Allocate National Target Sample Size to PSUs

```

*****
*   CREATE A DATA SET WITH THE CPI AREA POPULATIONS   *
*   INPUT: CE-ONLY PSU DEFINITIONS FILE FROM BLS      *
*****;

%MACRO LOADPSUS(NAME);
  DATA &NAME.;
    INFILE "T:\COMMON\CE Sampling Intervals\DATA\BLSFILES\&NAME..TXT"
LRECL=35 PAD MISSEVER;
    INPUT
      @1 REGION $1.
      @3 FIPSSST $2.
      @6 FIPSCY $3.
      @10 BLSPSU2K $4.
      @15 SR_NSR $1.
      @17 STRATPOP 8.0
      @26 UPROB 10.8;
    LENGTH CPI_AREA $4.;
    IF PUT(BLSPSU2K,$1.)='A' THEN CPI_AREA=BLSPSU2K;
    ELSE CPI_AREA = PUT(BLSPSU2K,$2.)||'00';
    PROC APPEND BASE=BLS_CE_FILE DATA=&NAME.;
    RUN;
%MEND;

%LOADPSUS(CENSOUT2000CPI);
%LOADPSUS(CENSOUT2000CE);

/* COLLAPSE COUNTY-LEVEL DATA SET TO PSU-LEVEL DATA SET */

PROC SORT DATA=BLS_CE_FILE NODUPKEY
  OUT=PSUS(KEEP=CPI_AREA BLSPSU2K STRATPOP);
  BY BLSPSU2K;
RUN;

PROC SUMMARY DATA=PSUS(WHERE=(CPI_AREA < 'Z100')) NWAY;
  CLASS CPI_AREA;
  VAR STRATPOP;
  OUTPUT OUT=CPI_AREAS(KEEP=CPI_AREA STRATPOP) SUM=;
DATA CPI_AREAS;
  SET CPI_AREAS;
  I+1;
DATA POP_DATA;
  ARRAY POP[36];
  DO UNTIL(LASTOBS);
    SET CPI_AREAS END=LASTOBS;
    POP[I]=STRATPOP;
  END;
  KEEP POP1-POP36;
RUN;

```

Appendix 7: SAS Program to Allocate National Target Sample Size to PSUs

```
* COMPUTE THE SQUARED DIFFERENCE BETWEEN EACH *
* CPI AREA'S PROPORTION OF THE POPULATION & ITS *
* PROPORTION OF THE SAMPLE. *
*****;
```

```
%MACRO MAC1;
SUM_POP = SUM(OF POP1-POP36);
%DO I=1 %TO 36;
    SQR&I = ((A&I/7300) - (POP&I/SUM_POP))**2;
%END;
%MEND MAC1;
```

```
*****
* SOLVE A CONSTRAINED LEAST SQUARES PROBLEM TO *
* FIND THE NUMBER OF HOUSING UNITS IN EACH PSU *
* THAT MINIMIZES THE SUM OF SQUARED DIFFERENCES *
*****;
```

```
PROC NLP DATA=POP_DATA NOPRINT
    OUT=RESULTS(KEEP=A1-A36)
```

```
/* CONVERGENCE CRITERIA */
    GCONV=1E-15
    GCONV2=1E-15
    ABSGCONV=1E-15
    FCONV2=1E-15
    MAXITER=100000 ;
```

```
/* DECISION VARIABLES */
    DECVAR A1-A36;
```

```
/* COMPUTE THE SQUARED DIFFERENCES */
    %MAC1;
```

```
/* SUM THE SQUARED DIFFERENCES */
    F1=SUM(OF SQR1-SQR36);
```

```
/* FUNCTION TO BE MINIMIZED */
    MIN F1;
```

```
/* PROBLEM CONSTRAINTS */
    BOUNDS A1-A36>=80;
    NLINCON F2=7300;
    F2=SUM(OF A1-A36);
```

```
RUN;
```

```
*****
* RE-LINK TO CPI-AREA CODES *
*****;
```

```
DATA RESULTS;
    ARRAY A[36] A1-A36;
    SET RESULTS;
    DO I = 1 TO 36;
```

Appendix 7: SAS Program to Allocate National Target Sample Size to PSUs

```

        ALLOCATION = A[I];
        OUTPUT;
    END;
    KEEP I ALLOCATION;
PROC SORT DATA=RESULTS; BY I;
PROC SORT DATA=CPI_AREAS; BY I;
DATA FINAL_NLP_ALLOCATION;
    MERGE CPI_AREAS RESULTS;
    BY I;
    DROP I;
RUN;

*****
*   PROPORTIONALLY ALLOCATE TARGET SAMPLE SIZES           *
*   TO PSUs WITHIN X AND Y CPI AREAS BY STRATUM POPS     *
*****;

/* ALLOCATE WITHIN CPI AREAS */

%MACRO ALLOCPSU(CPIAREA);

    DATA _NULL_;
        SET FINAL_NLP_ALLOCATION;
        WHERE CPI_AREA = "&CPIAREA.";
        CALL SYMPUT("CPIALLOC",ALLOCATION);
    RUN;
    DATA &CPIAREA.;
        SET PSUS;
        WHERE CPI_AREA = "&CPIAREA.";
        KEEP CPI_AREA BLSPSU2K STRATPOP;
    PROC FREQ DATA=&CPIAREA.;
        WEIGHT STRATPOP;
        TABLES BLSPSU2K /NOPRINT OUT=TEMP(DROP=COUNT);
    PROC SORT DATA=TEMP; BY BLSPSU2K;
    PROC SORT DATA=&CPIAREA.; BY BLSPSU2K;
    DATA &CPIAREA.;
        MERGE &CPIAREA. TEMP END=LASTONE;
        BY BLSPSU2K;
        PSU_ALLOCATION = &CPIALLOC. * PERCENT / 100 ;
        KEEP CPI_AREA BLSPSU2K PSU_ALLOCATION;
    RUN;

/* APPEND CPI AREA DATA SET TO CUMULATIVE DATA SET OF ALL PSUS */

PROC APPEND BASE=PSU_ALLOCATIONS DATA=&CPIAREA.;
RUN;

%MEND;

%ALLOCPSU(X100)
%ALLOCPSU(X200)
%ALLOCPSU(X300)
%ALLOCPSU(X400)
%ALLOCPSU(Y100)

```


Appendix 7: SAS Program to Allocate National Target Sample Size to PSUs

```

%ALLOCPSU(Y200)
%ALLOCPSU(Y300)
%ALLOCPSU(Y400);

*****
* APPEND "A" PSUs TO CUMULATIVE DATA SET OF ALL PSUS *
*****;

PROC SORT DATA=PSU_ALLOCATIONS;
  BY CPI_AREA;
PROC SORT DATA=FINAL_NLP_ALLOCATION;
  BY CPI_AREA;
DATA PSU_ALLOCATIONS;
  MERGE PSU_ALLOCATIONS(IN=XY) FINAL_NLP_ALLOCATION;
  BY CPI_AREA;
  IF NOT XY THEN DO;
    BLSPSU2K = CPI_AREA;
    PSU_ALLOCATION = ALLOCATION;
  END;
  RENAME ALLOCATION=CPI_AREA_ALLOCATION;
RUN;

*****
* PROPORTIONALLY ALLOCATE 400 UNITS AMONG Z PSUS *
* AND APPEND Z PSU ALLOCATION DATA SET *
*****;

PROC SORT DATA=BLS_CE_FILE(WHERE=(PUT(BLSPSU2K,$1.) = 'Z'))
  OUT=ZPSUS(KEEP=BLSPSU2K STRATPOP)
  NODUPKEY;
  BY BLSPSU2K;
PROC SUMMARY DATA=ZPSUS NWAY;
  VAR STRATPOP;
  OUTPUT OUT=ZSUM(KEEP=ZSUM) SUM=ZSUM;
DATA ZPSUS;
  SET ZSUM;
  DO UNTIL(LAST);
    SET ZPSUS END=LAST;
    PSU_ALLOCATION = 400 * ( STRATPOP / ZSUM );
    CPI_AREA = 'ZALL';
    CPI_AREA_ALLOCATION = 400;
  OUTPUT;
  END;
  KEEP CPI_AREA BLSPSU2K PSU_ALLOCATION CPI_AREA_ALLOCATION STRATPOP;

PROC APPEND BASE=PSU_ALLOCATIONS DATA=ZPSUS;
RUN;

*****
* DISPLAY PSU ALLOCATIONS AND COMPARE PSU ALLOCATION SUMS *
* WITHIN EACH CPI AREA WITH THE ORIGINAL CPI AREA ALLOCATION. *
*****;

```

Appendix 7: SAS Program to Allocate National Target Sample Size to PSUs

```
PROC SORT DATA=PSU_ALLOCATIONS;
  BY CPI_AREA BLSPSU2K;
DATA CPI_AREAS;
  SET PSU_ALLOCATIONS;
  BY CPI_AREA;
  IF FIRST.CPI_AREA AND CPI_AREA < 'Z100';
  KEEP CPI_AREA CPI_AREA_ALLOCATION;
RUN;
  TITLE 'CE REDESIGN 2000';
  TITLE2 'TARGET SAMPLE SIZE';
PROC PRINT DATA=CPI_AREAS NOOBS;
  TITLE3 'ALLOCATIONS FOR CPI AREAS';
  VAR CPI_AREA CPI_AREA_ALLOCATION;
  SUM CPI_AREA_ALLOCATION;
PROC PRINT DATA=PSU_ALLOCATIONS NOOBS;
  TITLE3 'ALLOCATIONS FOR X- AND Y-SIZE PSUS';
  WHERE PUT(CPI_AREA,$1.) IN ('X','Y');
  BY CPI_AREA;
  VAR BLSPSU2K PSU_ALLOCATION;
  SUM PSU_ALLOCATION;
  SUMBY CPI_AREA;
RUN;
PROC PRINT DATA=PSU_ALLOCATIONS NOOBS;
  TITLE3 'ALLOCATIONS FOR Z-SIZE PSUS';
  WHERE PUT(CPI_AREA,$1.) = 'Z';
  VAR BLSPSU2K PSU_ALLOCATION;
  SUM PSU_ALLOCATION;
RUN;
```

Appendix 8: SAS Program to Calculate PSU Inflation Factors

```

*****
*   USE CEQ AND CED INTERVIEW STATUS DATA FROM THE PERIOD   *
*   1999 - 2001 TO CALCULATE PARTICIPATION RATES FOR CPI AREAS *
*   AND ALSO NATIONAL RATES FOR EACH SURVEY.  FOR EACH CPI   *
*   AREA, CALCULATE A FACTOR WHICH IS A WEIGHTED AVERAGE    *
*   OF THE CPI AREA RATE AND THE NATIONAL RATE, WITH THE    *
*   CPI AREA RATE WEIGHTED 75% AND THE NATIONAL RATE        *
*   WEIGHTED 25%.                                           *
*****;

LIBNAME CEQ 'T:\COMMON\CE Sampling Intervals\DATA\CE DATA 99_01\CEQ';
LIBNAME CED 'T:\COMMON\CE Sampling Intervals\DATA\CE DATA 99_01\CED';

/* LOAD CEQ DATA */

%MACRO LOADCEQ(MONTH);
  DATA TEMP;
    LENGTH ID $9. STATUS $2.;
    ARRAY ISTAT[5] $ INTSTAT1-INTSTAT5;
    SET CEQ.INT&MONTH.;
    ID = PUT(CENSID,$9.);
    STATUS = ISTAT[INPUT(INTERI,1.)];
    IF STATUS = '01' THEN STATUS = 'I';
    ELSE STATUS = 'NI';
    KEEP ID STATUS;
  PROC APPEND DATA=TEMP BASE=CEQ;
  RUN;
%MEND;

%MACRO DOQYEAR(Y);
  %DO M = 1 %TO 9;
    %LOADCEQ(&Y.0&M.);
  %END;
  %DO M=10 %TO 12;
    %LOADCEQ(&Y.&M.);
  %END;
%MEND;

%DOQYEAR(99)
%DOQYEAR(00)
%DOQYEAR(01);

PROC SORT DATA=CEQ;
  BY ID;
RUN;
DATA IDTOCPA;
  INFILE 'T:\COMMON\CE Sampling Intervals\DATA\CE DATA 99_01\CEQ\
CE_CENSID_TO_CPI_AREA.TXT';
  INPUT @1 ID $9. @11 CPI_AREA $4.;
RUN;
PROC SORT;

```

Appendix 8: SAS Program to Calculate PSU Inflation Factors

```

BY ID;
RUN;
DATA CEQ;
  MERGE CEQ(IN=OK) IDTOCPA;
  BY ID;
  IF OK;

  /* CONVERT OBSERVATIONS FROM A212, A213, A214 TO CPI AREA X200 */
  IF CPI_AREA IN ('A212','A213','A214') THEN CPI_AREA = 'X200';

KEEP CPI_AREA STATUS;
RUN;

/* LOAD CED DATA */

%MACRO LOADCED(MONTH);
  DATA TEMP;
    LENGTH CPI_AREA $4. STATUS $2.;
    SET CED.CED_&MONTH.;
    SELECT(PUT(BLSPSU,$1.));
      WHEN('A') CPI_AREA=BLSPSU;
      WHEN('B') CPI_AREA='X' || SUBSTR(BLSPSU,2,1) || '00';
      WHEN('C') CPI_AREA='Y' || SUBSTR(BLSPSU,2,1) || '00';
      WHEN('D') CPI_AREA='Z' || SUBSTR(BLSPSU,2,1) || '00';
      OTHERWISE;
    END;

    /* CONVERT OBSERVATIONS FROM A212, A213, A214 TO CPI AREA X200 */
    IF CPI_AREA IN ('A212','A213','A214') THEN CPI_AREA = 'X200';

    DO W=1 TO 2;
      IF W=1 THEN STATUS=INTSTAT1;
      ELSE STATUS=INTSTAT2;
      IF STATUS = '01' THEN STATUS = 'I';
      ELSE STATUS = 'NI';
      OUTPUT;
    END;
    KEEP CPI_AREA STATUS;
  RUN;
  PROC APPEND DATA=TEMP BASE=CED;
  RUN;
%MEND;

%MACRO DODYEAR(Y);
  %DO M = 1 %TO 9;
    %LOADCED(&Y.0&M.);
  %END;
  %DO M=10 %TO 12;
    %LOADCED(&Y.&M.);
  %END;
%MEND;

%DODYEAR(99)

```

Appendix 8: SAS Program to Calculate PSU Inflation Factors

```

%DODYEAR(00)
%DODYEAR(01);

/* GET PARTICIPATION RATES AND CALCULATE FACTORS FOR EACH SURVEY */

%MACRO RATES(DSNAME);

  /* CPI AREA RATES */
  PROC SORT DATA=&DSNAME.;
    BY CPI_AREA;
  PROC FREQ DATA=&DSNAME.;
    BY CPI_AREA;
    TABLES STATUS /NOPRINT OUT=&DSNAME._CPI_AREA_RATES(DROP=COUNT);
  RUN;
  DATA &DSNAME._CPI_AREA_RATES;
    SET &DSNAME._CPI_AREA_RATES;
    WHERE STATUS='I';
    &DSNAME._CPI_AREA_RATE = PERCENT / 100;
    KEEP CPI_AREA &DSNAME._CPI_AREA_RATE;
  RUN;

  /* NATIONAL RATE */
  PROC FREQ DATA=&DSNAME.;
    TABLES STATUS /NOPRINT OUT=&DSNAME._NAT_RATE(DROP=COUNT);
  RUN;
  DATA &DSNAME._NAT_RATE;
    SET &DSNAME._NAT_RATE;
    WHERE STATUS='I';
    &DSNAME._NAT_RATE = PERCENT / 100;
    KEEP &DSNAME._NAT_RATE;
  RUN;

  /* CALCULATE CPI AREA FACTORS */
  DATA &DSNAME._FACTORS;
    SET &DSNAME._NAT_RATE;
    DO UNTIL(LAST);
      SET &DSNAME._CPI_AREA_RATES END=LAST;
      &DSNAME._CPIA_FACTOR =
        ( (0.75 * &DSNAME._CPI_AREA_RATE) + (0.25 * &DSNAME._NAT_RATE)
    );
    OUTPUT;
  END;
  KEEP CPI_AREA &DSNAME._CPIA_FACTOR &DSNAME._CPI_AREA_RATE
&DSNAME._NAT_RATE;
  RUN;
%MEND;

%RATES(CEQ);
%RATES(CED);

/* COMPARE THE TWO SURVEY FACTORS IN EACH CPI_AREA. THE LOWER FACTOR WILL */
/* BE USED TO INFLATE THE TARGET SAMPLE SIZES IN THE PSUS TO DETERMINE THE */
/* DESIGNATED SAMPLE SIZES FOR INITIAL SAMPLING. IF CED (DIARY) HAS THE */

```

Appendix 8: SAS Program to Calculate PSU Inflation Factors

```

/* LOWER FACTOR, THEN THE RATIO OF THE CEQ (INTERVIEW) FACTOR TO THE CED */
/* FACTOR WILL BE USED AS A SUBSAMPLING TAKE-EVERY TO REDUCE THE CEQ */
/* DESIGNATED SAMPLE SIZE AFTER INITIAL SAMPLING, AFTER THE TWO SAMPLES */
/* ARE SPLIT. */

PROC SORT DATA = CEQ_FACTORS;
  BY CPI_AREA;
PROC SORT DATA = CED_FACTORS;
  BY CPI_AREA;
DATA CEFACS;
  MERGE CEQ_FACTORS CED_FACTORS;
  BY CPI_AREA;
  CE_FACTOR = 1 / MIN( CEQ_CPIA_FACTOR, CED_CPIA_FACTOR);
  IF CED_CPIA_FACTOR < CEQ_CPIA_FACTOR THEN
    CEQ_TE = CEQ_CPIA_FACTOR / CED_CPIA_FACTOR;
  ELSE CEQ_TE = 1;
RUN;

/* BECAUSE THERE ARE NO 1990 PSUS CORRESPONDING TO THE Y100 CPI AREA */
/* EDIT THE DATA SET TO COPY THE X100 VALUES TO Y100. */

PROC SORT DATA=CEFACS;
  BY CPI_AREA;
DATA CEFACS;
  SET CEFACS;
  IF CPI_AREA = 'X100' THEN DO;
    OUTPUT;
    CPI_AREA='Y100';
    OUTPUT;
  END;
  ELSE OUTPUT;
PROC SORT; BY CPI_AREA;
RUN;

/* VIEW THE PARTICIPATION RATES AND INFLATION FACTORS */

PROC PRINT DATA=CEFACS LABEL NOOBS;
  TITLE 'CE REDESIGN 2000';
  TITLE2 'PARTICIPATION RATES AND INFLATION FACTORS';
  TITLE3 'BY REGION/SIZE CLASS';
  VAR CPI_AREA CEQ_CPI_AREA_RATE CEQ_NAT_RATE CEQ_CPIA_FACTOR
      CED_CPI_AREA_RATE CED_NAT_RATE CED_CPIA_FACTOR
      CE_FACTOR CEQ_TE;
  LABEL
    CPI_AREA='PSU GROUP'
    CEQ_CPI_AREA_RATE='CEQ PARTICIPATION RATE'
    CEQ_NAT_RATE='CEQ NATIONAL PARTICIPATION RATE'
    CEQ_CPIA_FACTOR='CEQ WEIGHTED AVERAGE RATE'
    CED_CPI_AREA_RATE='CED PARTICIPATION RATE'
    CED_NAT_RATE='CED NATIONAL PARTICIPATION RATE'
    CED_CPIA_FACTOR='CED WEIGHTED AVERAGE RATE'
    CE_FACTOR='INFLATION FACTOR USED'
    CEQ_TE='CEQ SUMSAMPLING TAKE-EVERY';
RUN;

```

Appendix 9: SAS Program to Calculate PSU Designated Sample Sizes

```
*****
*   CALCULATE CE DESIGNATED SAMPLE SIZES TO BE USED FOR   *
*   INITIAL SAMPLING.  DIVIDE TARGET SAMPLE ALLOCATED   *
*   TO EACH PSU BY THE CORRESPONDING CPI AREA FACTOR   *
*   CALCULATED FROM CE 1999-2001 RESPONSE RATES.       *
*****;

* Note: The allocation program and the inflation factor program must be run
before this program;

DATA PSU_ALLOCATIONS;
  SET PSU_ALLOCATIONS;
  IF CPI_AREA='ZALL' THEN CPI_AREA=PUT(BLSPSU2K,$2.)||'00';
PROC SORT DATA=PSU_ALLOCATIONS;
  BY CPI_AREA;
RUN;

PROC SORT DATA=CEFACS;
  BY CPI_AREA;
RUN;

/* MERGE DATA SETS AND CALCULATE DESIGNATED SAMPLE SIZES */

DATA CE_PSU_DSS;
  MERGE PSU_ALLOCATIONS CEFACS;
  BY CPI_AREA;

  /* MULTIPLY BY 2 BECAUSE TWO SURVEY SAMPLES NEEDED, CEQ AND CED */
  PSU_DSS = 2 * PSU_ALLOCATION * CE_FACTOR ;

  KEEP BLSPSU2K PSU_ALLOCATION CE_FACTOR PSU_DSS;
RUN;

/* DISPLAY PSU DESIGNATED SAMPLE SIZES AND TOTAL DESIGNATED SAMPLE SIZE */

PROC PRINT DATA=CE_PSU_DSS LABEL NOOBS;
  TITLE 'CE 2000 REDESIGN';
  TITLE2 'DESIGNATED SAMPLE SIZES';
  VAR BLSPSU2K PSU_ALLOCATION CE_FACTOR PSU_DSS;
  SUM PSU_DSS;
  LABEL
    BLSPSU2K = 'STRAT PSU'
    PSU_ALLOCATION = 'ALLOCATED TARGET SAMPLE SIZE'
    CE_FACTOR = 'NON-RESPONSE INFLATION FACTOR'
    PSU_DSS = 'DESIGNATED SAMPLE SIZE';
RUN;
```

Appendix 10: SAS Program to Calculate PSU Sampling Intervals

```
*****
*   CALCULATE CE WITHIN-PSU SAMPLING INTERVALS.  SAMPLING INTERVAL   *
*   WILL BE THE RATIO OF THE PSU MEASURE OF SIZE (2005 PROJECTED # OF *
*   HOUSING UNITS) TO THE DESIGNATED SAMPLE SIZE.                      *
*****;
```

```
LIBNAME CENSUS2K 'T:\COMMON\CE Sampling Intervals\DATA\CENSUS DATA';
```

* Note: The Allocation, Rates, and Designated Sample Size programs must be run before this one. ;

```
/* GET PROJECTED 2005 HOUSING UNIT COUNTS BY COUNTY */
```

```
DATA PROJ_HU_CTS;
  SET CENSUS2K.Proj_05_hu_counts_by_cty;
  RENAME STATE=FIPSST
         COUNTY=FIPSCTY
         PHU05ACSNU=HU_CT_PROJ;
  KEEP STATE COUNTY PHU05ACSNU;
RUN;
```

```
/* MODIFY TO CORRECT FOR PROJECTIONS IN NORTH DAKOTA AND WEST VIRGINIA */
/* WHICH WERE LESS THAN THE CENSUS 2000 COUNTS FOR THOSE STATES, AND */
/* ALSO MODIFY THE DC PROJECTION, WHICH IS DEEMED UNREALISTIC.  THE */
/* NORTH DAKOTA AND WEST VIRGINIA PROJECTIONS WILL BE REPLACED BY */
/* THE CENSUS 2000 COUNTS, AND THE DC PROJECTION WILL BE REPLACED BY */
/* A HOUSING UNIT ESTIMATE OF 268,504 WHICH IS THE ESTIMATE BEING USED */
/* BY CPS AND SIPP FOR DC.                                          */
```

```
DATA ND_WV_2000_HUS;
  SET CENSUS2K.C2KCOUNT;
  WHERE FIPSST IN ('38', '54');
  KEEP FIPSST FIPSCTY CENSUS2000HOUSINGUNITCOUNT;
PROC SORT;
  BY FIPSST FIPSCTY;
PROC SORT DATA=PROJ_HU_CTS;
  BY FIPSST FIPSCTY;
DATA PROJ_HU_CTS;
  MERGE PROJ_HU_CTS(IN=P) ND_WV_2000_HUS(IN=C);
  BY FIPSST FIPSCTY;
  IF P AND C THEN HU_CT_PROJ = CENSUS2000HOUSINGUNITCOUNT;
  IF FIPSST='11' THEN HU_CT_PROJ = 268504;
  KEEP FIPSST FIPSCTY HU_CT_PROJ;
RUN;
```

```
/* APPEND PROJECTED 2005 HU COUNTS TO CE PSU FILE */
```

```
PROC SORT DATA=BLS_CE_FILE;
  BY FIPSST FIPSCTY;
PROC SORT DATA=PROJ_HU_CTS;
  BY FIPSST FIPSCTY;
DATA BLS_CE_FILE;
  MERGE BLS_CE_FILE(IN=OK) PROJ_HU_CTS;
  BY FIPSST FIPSCTY;
```


Appendix 10: SAS Program to Calculate PSU Sampling Intervals

```
IF OK;
RUN;

/* GET PSU MEASURE OF SIZE */

PROC SUMMARY DATA=BLS_CE_FILE NWAY;
  CLASS BLSPSU2K;
  VAR HU_CT_PROJ;
  OUTPUT OUT=PSUHUCTS(KEEP=BLSPSU2K HU_CT_PROJ) SUM=;
RUN;

/* MERGE DATA SETS AND CALCULATE SAMPLING INTERVALS */

PROC SORT DATA=PSUHUCTS;
  BY BLSPSU2K;
PROC SORT DATA=CE_PSU_DSS;
  BY BLSPSU2K;
DATA SAMPINTS;
  MERGE PSUHUCTS CE_PSU_DSS;
  BY BLSPSU2K;
  SAMPINT = HU_CT_PROJ / PSU_DSS ;
RUN;

/* VIEW FINAL DATA SET */

PROC PRINT DATA=SAMPINTS LABEL NOOBS;
  TITLE 'CE 2000 REDESIGN';
  TITLE2 'WITHIN-PSU SAMPLING INTERVALS';
  VAR BLSPSU2K HU_CT_PROJ PSU_DSS SAMPINT;
  FORMAT SAMPINT COMMA14.4;
  LABEL
    BLSPSU2K = 'STRAT PSU'
    HU_CT_PROJ = 'PROJECTED 2005 HU COUNT'
    PSU_DSS = 'DESIGNATED SAMPLE SIZE'
    SAMPINT = 'PSU SAMPLING INTERVAL';
RUN;
```

Appendix 10: SAS Program to Calculate PSU Sampling Intervals

```
*****
* PROJECT 2005 PERMIT COUNTS BY COUNTY BASED ON FILES FROM MCD WHICH *
* DSMD RECEIVED FOR THE YEARS 1997 THROUGH 2001 AND USED TO BUILD *
* THEIR 1990-BASED DESIGN PERMIT DATA UNIVERSE FOR NEW CONSTRUCTION *
* SAMPLING. FOR EACH COUNTY, THE PROJECTION WILL BE THE COUNT VALUE *
* OF THE POINT ON THE LEAST SQUARES REGRESSION LINE CORRESPONDING TO *
* THE YEAR 2005. *
*****;
```

* Note: The CE Sampling Interval Programs must be run before this one.;

```
DATA PROJECTED2005PERMITS;
  SET census2k.PERMITBYCTY;
  ARRAY YR_[5];
  ARRAY CT[5] COUNT1997-COUNT2001;
  ARRAY RESIDUAL[5];
  RETAIN YR_1-YR_5 (1997 1998 1999 2000 2001);
  YR_SUM = SUM(OF YR_[*]);
  CTSUM = SUM(OF CT[*]);
  YR_SQSUM = 0; YR_CTSUM = 0;
  DO I = 1 TO 5;
    YR_SQSUM + YR_[I]**2;
    YR_CTSUM + (YR_[I]*CT[I]);
  END;
  SLOPE = ( ( 5 * YR_CTSUM ) - ( YR_SUM * CTSUM ) )
          /
          ( ( 5 * YR_SQSUM ) - YR_SUM**2 );
  INTERCEPT = ( CTSUM - ( SLOPE * YR_SUM ) )
                /
                5;
  DO I = 1 TO 5;
    RESIDUAL[I] = ABS( CT[I] - ( (SLOPE * YR_[I]) + INTERCEPT ) );
  END;

  PROJECTED2005COUNT = CEIL((2005 * SLOPE) + INTERCEPT);
  IF PROJECTED2005COUNT > 0 THEN
    RESIDUALRATIO = MEAN(OF RESIDUAL[*]) / PROJECTED2005COUNT ;
  ELSE RESIDUALRATIO = 2;
  IF SLOPE < 0 OR RESIDUALRATIO > 1 THEN DO;
    ORIGINAL_PROJECTION = PROJECTED2005COUNT;
    PROJECTED2005COUNT = CEIL(MEAN(OF CT[*]));
    MEAN_USED = 1;
  END;
  ELSE MEAN_USED=0;

  /* RECODE FIPS COUNTY FOR MIAMI-DADE, FLORIDA */
  IF FIPSST = '12' AND FIPSCTY='025' THEN FIPSCTY='086';

  OUTPUT;
RUN;
```

```
*****
* SUBSET CPI COUNTIES AND GET SUM ACROSS ALL CPI PSUS *
*****;
```

Appendix 10: SAS Program to Calculate PSU Sampling Intervals

```

DATA CPICTYS;
  INFILE
    'T:\COMMON\CE Sampling Intervals\DATA\BLSFILES\CENSOUT2000CPI.TXT'
  MISSOVER;
  INPUT @3 FIPSST $2. @6 FIPSCTY $3.;
  KEEP FIPSST FIPSCTY;
PROC SORT; BY FIPSST FIPSCTY;
PROC SORT DATA=PROJECTED2005PERMITS;
  BY FIPSST FIPSCTY;
DATA CPIPMTCTS;
  MERGE CPICTYS(IN=CPI) PROJECTED2005PERMITS;
  BY FIPSST FIPSCTY;
  IF CPI;
PROC SUMMARY DATA=CPIMPMTCTS;
  OUTPUT OUT=CPIMPMTSUM(KEEP=NAT2005PP) SUM=NAT2005PP;
  VAR PROJECTED2005COUNT;
RUN;

DATA CPI_SAMPINT;
  SET CPIMPMTSUM;

  /* SAMPLING INTERVAL IS

      (PROJECTED NUMBER OF PERMITS IN 2005 IN CPI-U SAMPLE COUNTIES) X
4
  -----
                        1440

  BECAUSE ANNUAL SAMPLE SHOULD BE 1440 PERMIT ADDRESSES, AND WE EXPECT A
  CLUSTER OF
  4 ADDRESSES FOR EACH HIT
  */

  CPISAMPINT = (NAT2005PP / 1440) * 4;
RUN;

OPTIONS NODATE NONUMBER NOCENTER LS=97 PS=51;

*****
*   DISPLAY THE NATIONAL SAMPLING INTERVAL   *
*****;

PROC PRINT DATA=CPI_SAMPINT NOOBS LABEL;
  TITLE 'THE CPI PERMIT NEW CONSTRUCTION HOUSING SAMPLE';
  TITLE2 'NATIONAL SAMPLING INTERVAL FOR THE CENSUS-2000 BASED DESIGN';
  VAR CPISAMPINT;
  LABEL CPISAMPINT='NATIONAL SAMPLING INTERVAL';
  FORMAT CPISAMPINT COMMA10.4;
RUN;

*****
*   LIST HISTORICAL COUNTS AND 2005 PROJECTIONS FOR CPI COUNTIES   *

```

Appendix 10: SAS Program to Calculate PSU Sampling Intervals

```
*****;  
PROC PRINT DATA=CPIPMTCTS(KEEP=COUNT1997-COUNT2001 PROJECTED2005COUNT FIPSST  
FIPSCTY) N LABEL;  
  TITLE 'PROJECTIONS OF PERMIT COUNTS';  
  TITLE2 'IN COUNTIES SELECTED FOR THE CENSUS 2000-BASED CPI SAMPLE DESIGN';  
  TITLE3 'BASED ON PERMIT COUNTS FROM THE YEARS 1997-2001';  
  ID FIPSST FIPSCTY;  
  VAR COUNT1997-COUNT2001 PROJECTED2005COUNT;  
  LABEL  
    FIPSST = 'FIPS STATE'  
    FIPSCTY = 'FIPS COUNTY'  
    COUNT1997 = '1997 COUNT'  
    COUNT1998 = '1998 COUNT'  
    COUNT1999 = '1999 COUNT'  
    COUNT2000 = '2000 COUNT'  
    COUNT2001 = '2001 COUNT'  
    PROJECTED2005COUNT = 'PROJECTED 2005 COUNT';  
  SUM _NUMERIC_;  
  FORMAT _NUMERIC_ COMMA10.0;  
RUN;
```