**date:** November 7, 2013

 **memorandum for:** W. JOHN LAYNG, Assistant Commissioner

 *Division of Consumer Prices and Price Indexes*

 *Office of Prices and Living Conditions*

 **memorandum from:** STEVE PABEN, Chief

 *Division of Price Statistical Methods*

 *Office of Prices and Living Conditions*

 **subject:** Final Materials for the 2010 Decennial Area Sample Design

 Revised 11/07/13

The CPI-CE Area Redesign Team has completed its mission of determining the next geographic area sample design for the Consumer Expenditure Survey (CE) and for the CPI surveys (Housing, C&S, and TPOPS).[[1]](#footnote-1) The new area design is based upon the demographics of the United States as of the 2010 Decennial Census, and is hereby referred to as the ‘2010 Design’. The new design was approved by the BLS Acting Commissioner, John Galvin, in August 2012. This memo (i) provides a summary of the essential features of the 2010 Design, (ii) identifies critical tasks associated with converting the four surveys to the new design along with proposed milestone dates for these tasks, and (iii) contains the pertinent sample design materials necessary for DCPPI to complete its area sample rotation requirements and plans.

Due to data collection difficulties with the selected geographic area, Big Stone Gap VA, the non-self-representing PSU N35Q will be represented by Clarksburg WV. These two geographic areas have similar demographics, in that they come from the same stratum and are both micropolitan CBSAs. This revised memorandum updates the previous memorandum for this change in CPI PSUs. It should be noted that, CE will continue to survey Big Stone Gap VA but all other geographic PSU will be the same.

1. **Summary of 2010 Design Features**

The CPI-CE Area Redesign Team began work on the new area design in FY10Q1. The team issued a series of reports and recommendations on sampling design issues prior to determining the final set of selected PSUs for the 2010 sample. These reports contain comprehensive information on (i) PSU definition and stratification ([Issued September 2010](http://oplc.sp.bls.gov/SMD/CPICE%20Area%20Sample%20Team/Forms/AllItems.aspx)), (ii) the use of Census Divisions to define non self-representing areas ([Issued March 2011](http://oplc.sp.bls.gov/SMD/CPICE%20Area%20Sample%20Team/Recommended%20switch%20to%20divisions.docx)), (iii) overlap maximization ([Issued September 2011](http://oplc.sp.bls.gov/SMD/CPICE%20Area%20Sample%20Team/Overlap%20Maximization%20Recommendations/Sample%20Overlap%20and%20Sample%20Rotation%20Memo.docx)), and (iv) the recommended number of PSUs ([Issued February 2012](http://oplc.sp.bls.gov/SMD/CPICE%20Area%20Sample%20Team/Memo%20on%20number%20of%20PSUs%20for%20CPI-CE%20Management%20-%20final%202-2-12.docx)). All of these reports are available on the team’s SharePoint site at the hyperlinks included above. A summary of the approved essential features of the 2010 Design, as detailed in the linked reports, is provided here:

* Population Target:
	+ CE: total non-institutional population of the United States
	+ CPI: urban, non-institutional population of the United States
* Primary Sampling Unit (PSU) Definition:
	+ Urban (CE and CPI): core-based-statistical-area (CBSA)[[2]](#footnote-2)
	+ Rural (CE only): clusters of mostly contiguous non-CBSA counties
	+ NOTE: The use of CBSAs represents a change from the 1990 Design. Population living in a rural area of a county that is included in a micropolitan CBSA will now be considered urban. Some small urban clusters in non-CBSAs will be excluded. As a result of this definitional change, the population coverage of the CPI will expand by 4.2%
* Number of Selected PSUs:
	+ CPI: 75 (a reduction of 12 from the 87 in the 1990 Design); this would have eventually allowed both surveys to collect data from the same set of geographic areas, reducing measurement error
	+ CE: 91 (the same 75 urban PSUs as originally sampled by CPI plus an additional 16 rural PSUs); same number as in the 2000 Design[[3]](#footnote-3)
* Alaska and Hawaii:
	+ Alaska: the four CBSAs in Alaska (Anchorage, Fairbanks, Juneau, Ketchikan) are grouped into a state stratum; Anchorage is selected to represent the stratum with certainty
	+ Hawaii: the four CBSAs in Hawaii (Honolulu, Kahului, Hilo, Kapaa) are grouped into a state stratum; Honolulu is selected to represent the stratum with certainty
* Population Cut-off for Self-representation: 2.5 million
	+ Number of self-representing PSUs: 21 (down from 29 in the 1990 Design)
	+ Number of other certainty PSUs: 2 (Anchorage and Honolulu)
	+ Number of non-self-representing PSUs: 52
* Stratification Groups: Census Divisions
	+ Non self-representing CBSAs will be grouped by Census Divisions for the purposes of stratifying CBSAs and PSU selection; this represents a change from the 1990 Design, which grouped PSUs by Region X Size Class
* Number of Elementary CPI Areas: 32
	+ 21 self-representing PSUs + 2 certainty strata + 9 Census Divisions
	+ This represents a reduction of 6 areas from the 1990 Design (29 self-representing + 2 certainty + 7 region-size groups)
* Stratification within Census Division:
	+ Model based, grouping CBSAs together based on geographic proximity (latitude and longitude), mean household income, and mean property value
	+ CBSA characteristics determined using 2006-2010 American Community Survey (ACS) data
* PSU Selection Methods:
	+ Overlap Maximization used (Ernst Method)
	+ Controlled Selection used (to control for the appropriate number of micropolitan CBSA selections)
	+ CBSA probabilities of selection based upon 2010 Decennial Census

**Chart 1.** Summary and Map of 2010 Geographic Area Design





**B. Sample Rotation Activities, Milestone Tasks, and Key Dates**

For the 2010 Design, the CE will convert to the new sample of PSUs in its entirety in 2015. Logistically, the CE will begin conducting interviews in the new area design in February 2015 and cease data collection in dropping PSUs by March 2015. CE will support estimation on the new sample design by the first calendar quarter of 2015.

A continuous sample rotation process is planned for Housing, C&S, and TPOPS. Under this plan, new PSUs will be introduced into the sample of each survey and dropping PSUs will be eliminated from each survey, over a multi-year span. The benefit of a continuous sample rotation process is to distribute the cost of adding and dropping PSUs into Housing and C&S over a period of years, rather than incur a spike in data collection costs leading up to the point when the computation of the CPI converts to the new structure.

The CPI-CE Area Redesign Team was not charged with the task of recommending a sample rotation plan. The CPIMG will charter a new team for this work. However, the CPI-CE Area Redesign Team has identified the following list of sample rotation activities, constraints, and issues that will influence the continuous rotation plan for Housing, C&S, and TPOPS.

**Outgoing 1990 Design vs. Incoming 2010 Design**

Each PSU selected for the 2010 Design sample can be classified into one of eight groups based upon the sampling status of the PSU in the 1990 Design. These groupings form the foundation of the rotation requirements. For the purposes of labeling, let S=a self-representing PSU and let N=a non-self-representing PSU. The groups are:

1. **S continuing as S.** These are the **21** current ‘A-sized’ PSUs that are reselected as self-representing ‘S-sized’ PSUs along with 2 other ‘A-sized PSUs, Anchorage and Honolulu, that are re-selected with certainty to represent the CBSAs in Alaska and Hawaii respectively. These PSUs “overlap” the 1990 and 2010 geographic designs with no city size change and thus no further action is required. To reduce overall geographic sample conversion costs, outlet rotation could be suspended for a period of time in these PSUs, to accommodate the cost of bringing in new PSUs. Many continuing A-size PSUs may drop or add counties, as PSU definitions historically change with each revision. Depending on the change, this will either slightly expand or contract the geographic territory that is eligible for Housing and C&S data collection activities, since eligible housing units and outlets are restricted to those found inside each PSU definition. See below for a detailed list of these PSUs.
2. **S continuing as N.** These are the 5 current ‘A-sized’ PSUs that are reselected as non-self-representing PSUs. No change in Housing or C&S sample rotation scheduling is required. Underlying sample sizes in TPOPS, C&S, and Housing should be downsized to non-certainty levels. These PSUs lose their elementary index area status. Hence, indexes for these PSUs will not automatically be calculated and made available for publication, unless a decision is made to calculate and release them as special-relative-calculation indexes (SRCs).
3. **S dropped.** This is the 1 current ‘A-sized’ PSUs that is not reselected in the 2010 sample. This PSU may serve as a proxy for a new PSU, or it may be dropped at the earliest possible date. Published indexes for this PSU will terminate.
4. **N continuing as S**. There are none of these.
5. **N continuing as N**. These are the 26 current ‘B-sized’ and ‘C-sized’ PSUs that are reselected and continue as ‘N-sized’ PSUs. Though no further action is required, to balance the cost associated with cycling in new PSUs, outlet rotation may be suspended for a period of time. PSU definitions change, as noted, which may impact collection costs. Slight changes to Housing, TPOPS, or C&S sample sizes may result.
6. **N dropped**. These are the 30 ‘B-sized’ and ‘C-sized’ PSUs in the 1990 sample that are not reselected for the 2010 sample. These PSUs may serve as proxy PSUs for a new PSU or be dropped from TPOPS, C&S, and Housing at the earliest possible date.



1. **New S**. There are none of these.
2. **New N**. These are the 21 ‘N-sized’ PSUs that are not in the 1990 sample, but are selected as part of the 2010 sample. Housing, TPOPS and C&S operations will need to be scheduled in these PSUs. Under the continuous rotation scheme, a “proxy” PSU (from the S dropped or N dropped group) will be assigned to each new PSU; price data collected in the proxy PSU will be used to represent the new PSU, until the new PSU is rotated into Housing and C&S.



Proxy Assignments and Constraints All dropping PSUs are proxy candidates for new PSUs until the complete set of new PSUs is rotated in. The ideal proxy for a new PSU is a PSU from the same 2010 Design geographic stratum. Three situations exist:

* *A stratum with one and only one dropping PSU*. If a new PSU is in a stratum with one and only one PSU that is dropping, the old PSU automatically serves as the proxy for the new PSU until it is rotated in. These automatically matched PSUs are taken out of the pool of remaining potential proxy PSUs.

New PSU Dropping PSU(Proxies)

Flint MI Youngstown, OH

Huntsville, AL Picayune, MS

Janesville, WI Wausau, WI

Louisville, KY Morristown, TN

Paris, TX Beaumont, TX

Russellville, AR Pine Bluff, AR

St. George, UT Boise City, ID

Winston-Salem, NC Florence, SC

* *A stratum with more than one dropping PSU*. If a new PSU is in a stratum with more than one dropping PSU, then one of the dropping PSUs will serve as proxy while the others can either be dropped or serve as potential proxies for other PSUs.

New PSU Old PSU (Possible Proxies)

Charlotte, NC Statesboro, GA \*

 Albany, GA

Frankfort, IN Evansville, IN

 Decatur, IL \*

 Mount Vernon, IL

Moses Lake, WA Pullman, WA \*

 Bend, OR

Wahpeton, ND Brookings, SD \*

 Kansas City, MO

Note: \* the closest geographic proxy to new PSU

* *A stratum with no dropping PSU*. If a new PSU is in a stratum that contains no active PSU in the CPI sample, the PSU will need to be assigned a proxy from another stratum that contains multiple dropping PSUs, or if none are available, the PSU must be scheduled to rotate with the first rotation group, such that Housing and C&S data are collected in the new PSU and available for use in index estimation, prior to the date the CPI is officially computed on the new geographic structure.

Geographic Holes

Clarksburg, WV Orlando, FL

Jacksonville, NC Rochester, NY

Meridian, MS Santa Rosa, CA

Omaha, NE Wichita, KS

Wilmington, NC

The following rules and constraints will govern the proxy assignment and rotation scheduling:

**Proxy assignment:**

* A dropping PSU can be used as proxy for one and only one new PSU
* A continuing PSU cannot serve as proxy for another PSU
* Proxy assignment for each PSU in a stratum with exactly one dropping PSUs is made prior to the proxy assignment of each PSU in a stratum with no dropping PSUs
* Proxy assignment for each PSU in a stratum with multiple dropping PSUs is made prior to the proxy assignment of each PSU in a stratum with no dropping PSUs
* Proxy assignment for each PSU in a stratum with no dropping PSUs:
	+ The proxy for each PSU in a stratum with no dropping PSUs is the dropping PSU from the remaining set of eligible dropping PSUs that is the nearest neighbor
	+ The nearest neighbor can be one in a different Census Division or Census Region
	+ If a given proxy PSU is the nearest neighbor to more than one new PSU, it will be assigned as proxy to the new PSU that is closer in distance
	+ The nearest neighbor proxy PSU must be within 200 miles of the new PSU
	+ If no eligible proxy exists, the new PSU is considered a “geographic hole”

**Rotation Scheduling:**

The following rules reflect what the CRPT team decided upon for the 2000 Decennial area sample. It is possible that the numbers of PSUs that can be rotated per year has changed. OFO should review these rules.

* No more than six PSUs can be rotated into the sample per calendar year
* No more than two PSUs per BLS Regional Office may be rotated per calendar year
* Rotation priority of new PSUs will be in the following order:
	+ Geographic holes (PSUs with no eligible proxies)
	+ PSUs with nearest neighbor, non-stratum proxies
	+ PSUs with within-stratum proxies
* Rotation priority of dropping PSUs will be in the following order
	+ PSUs not serving as a proxy
	+ PSUs serving as proxy
		- Drop in sequence as determined by rotation schedule of corresponding new PSU
		- Within priority group, dropping PSUs with the longest period of time in the CPI have priority to drop sooner
* After proxy assignments are made, if there are more than six geographic holes, it may not be possible to rotate all of them into Housing and C&S prior to the date the CPI needs to compute indexes on the new structure; under this scenario, options include:
	+ Relaxing the 6-PSU-per-year constraint
	+ Expanding the nearest neighbor constraint to include proxies beyond 200 miles
	+ Imputing price-change for each new PSU until it can be rotated in, using corresponding elementary area price, by item (this option is disfavored).

**Rotation Milestone Tasks and Key Dates**

The table below identifies critical milestone activities in the sample rotation plan, and proposed milestone dates for each activity. The critical-path activities are those associated with the first group of new PSUs slated to be rotated into the sample. The dates below are based on the assumption that the 2010 Design geographic structure will be first used to calculate the January 2018 CPI-U and CPI-W indexes, published in February 2018.

**Table 5** List of Critical Geographic Revision Sample Rotation Tasks and Dates

|  |  |
| --- | --- |
| **GEOGRAPHIC REVISION TASK** | **DATE** |
| First rotation group of new PSUs |   |
| Begin TPOPS at Census | Oct 2015 |
| Item and outlet selection (SMS)  | Aug 2016 |
| Begin Initiation of housing sample | Jan 2017 |
| Begin C&S Initiation | Feb 2017 |
| Begin pricing of housing units | Jul 2017 |
| Begin pricing of C&S items | Sep 2017 |
| First use in index | Feb 2018 |

Remaining geographic sample rotation activities to be addressed and resolved by DCPPI include:

* Determine if urban population outside of CBSAs should be accounted for via weight adjustments
* Determine if the region/size class index series Should continue to be produced and published
* Determine if we should continue publication of self-representing PSUs continuing as non-self-representing PSUs
* Publication attributes of new index series (Census Divisions): base period, start month, etc.
* Reoptimize IOOP model (SMD)
* Bridging requirements for Housing
* Bridging requirements for C&S
* Determine shadow PSUs for Housing
* EWPS concordance requirements
* Make final proxy assignments
* Make final PSU rotation groupings
* Schedule rotation:
	+ - When to drop non proxies
		- When and if to suspend and resume TPOPS in continuing PSUS
		- When to convert TPOPS to new PSU definitions in continuing PSUs
		- When to begin first sample rotation of new PSUs after initial sample
		- Impact on CPI Housing Initiative (if and when to suspend replacement samples; whether or not to eliminate rotation in dropping PSUs serving as proxies, etc.)

1. **2010 Decennial Area Sample Design Information**

The attached Excel workbook contains all of the pertinent information defining the 2010 Decennial Area Sample Design. This information should be used in making final rotation plans and schedules. The workbook contains the following worksheets:

* **Data Dictionary.** Contains a list of every variable in the other worksheets, and a description of each variable.
* **County\_View**. One record for each of the 3,144 counties, or county-equivalents, in the United States. Exception: some counties in New England appear more than once, if minor-civil-divisions (i.e., townships) in the county were mapped to different PSUs in the 1990 area design. There will be a separate observation in the worksheet for the county and the 1990 PSU designation.
* **CBSA\_ View**. One record for each of 918 CBSAs in the United States, plus one record for each of the rural groupings
* **Sample\_View.** One record for each of the 91 selected CPI PSUs
* **Nearest Neighbor.** For each new PSU, lists the distance from the new PSU to each old PSU that is dropping.
* **2010 Area Design Summary.** Contains the list of selected PSUs, their county definitions and stratum and PSU populations.
1. CPI-CE Area Redesign Team members are Bob Eddy (OPLC); Steve Paben, Dave Swanson, Bill Johnson, John Schilp, and Susan King (SMD), Carolyn Pickering (CE); and Rob Cage (DCPPI). [↑](#footnote-ref-1)
2. A Core Based Statistical Area (CBSA) is a U.S. geographic area defined by the Office of Management and Budget (OMB) for use by Federal statistical agencies in collecting, tabulating, and publishing Federal statistics. A CBSA is a collective term for both metropolitan (metro) and micropolitan (micro) areas. A metro area contains a core urban area of 50,000 or more population, and a micro area contains an urban core of at least 10,000 (but less than 50,000) population. Each metro or micro area consists of one or more counties and includes the counties containing the core urban area, as well as any adjacent counties that have a high degree of social and economic integration (as measured by commuting to work) with the urban core. [↑](#footnote-ref-2)
3. Due to the revised CPI sample, Clarksburg WV will replace Big Stone Gap VA for CPI data collection . Consumer Expenditure survey will continue with collection efforts in originally sampled PSUs, which include Big Stone Gap VA. This takes the CPI and CE out of synch with one non-self-represeting area, N35Q. [↑](#footnote-ref-3)